



09-22-00

A

Box 5241
PTO/SB/05 (2/98)

UTILITY

PATENT APPLICATION
TRANSMITTAL

Attorney Docket No.

210121.475C7

First Inventor or Application Identifier

Steven G. Reed

Title

COMPOSITIONS AND METHODS FOR THE THERAPY
AND DIAGNOSIS OF LUNG CANCER

Express Mail Label No.

EL615232339US

Only for nonprovisional applications under 37 CFR § 1.53(b)

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

ADDRESS TO:

Box Patent Application
Assistant Commissioner for Patent
Washington, D.C. 202311. ☐ General Authorization Form & Fee Transmittal
(Submit an original and a duplicate for fee processing)2. ☒ Specification [Total Pages] **137**
(preferred arrangement set forth below)

- Descriptive Title of the Invention
- Cross References to Related Applications
- Statement Regarding Fed sponsored R & D
- Reference to Microfiche Appendix
- Background of the Invention

- Brief Summary of the Invention
- Brief Description of the Drawings (if filed)
- Detailed Description
- Claim(s)
- Abstract of the Disclosure

☐ Drawing(s) (35 USC 113) [Total Sheets] Oath or Declaration [Total Pages] a. ☐ Newly executed (original or copy)b. ☐ Copy from a prior application (37 CFR 1.63(d))
(for continuation/divisional with Box 17 completed)i. ☐ **DELETION OF INVENTOR(S)**
Signed statement attached deleting
inventor(s) named in the prior application,
see 37 CFR 1.63(d)(2) and 1.33(b)5. ☐ Incorporation By Reference (useable if box 4b is
checked) The entire disclosure of the prior application,
from which a copy of the oath or declaration is supplied
under Box 4b, is considered to be part of the disclosure of
the accompanying application and is hereby incorporated
by reference therein.6. ☐ Microfiche Computer Program (Appendix)7. Nucleotide and Amino Acid Sequence Submission
(if applicable, all necessary)

- a. ☒ Computer-Readable Copy
- b. ☒ Paper Copy (identical to computer copy)
- c. ☒ Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

8. ☐ Assignment Papers (cover sheet & document(s))9. ☐ 37 CFR 3.73(b) Statement ☐ Power of Attorney
(when there is an assignee)10. ☐ English Translation Document (if applicable)11. ☒ Information Disclosure
Statement (IDS)/PTO-1449 ☒ Copies of IDS
Citations12. ☐ Preliminary Amendment13. ☒ Return Receipt Postcard14. ☐ Small Entity
Statement(s) ☐ Statement filed in prior application,
Status still proper and desired15. ☐ Certified Copy of Priority Document(s)
(if foreign priority is claimed)16. ☒ Other: Certificate of Express Mail

17. If a CONTINUING APPLICATION, check appropriate box and supply the requisite information below and in a preliminary amendment

☐ Continuation ☐ Divisional ☒ Continuation-In-Part (CIP) of prior Application No.: **09/640,878**Prior application information: Examiner **not assigned** Group / Art Unit **not assigned**☐ Claims the benefit of Provisional Application No.

CORRESPONDENCE ADDRESS

Jane E. R. Potter
Seed Intellectual Property Law Group PLLC
701 Fifth Avenue, Suite 6300
Seattle, Washington 98104-7092
Phone: (206) 622-4900 Fax: (206) 682-6031

Respectfully submitted,

TYPED or PRINTED NAME Jane E. R. PotterSIGNATURE Jane E. R. Potter

u:\sharons\210121\475C7

REGISTRATION NO. 33,332Date September 20, 2000

jc541 U.S. PTO
60/234517

09/22/00

u:\sharons\210121\475

COMPOSITIONS AND METHODS FOR THE THERAPY
AND DIAGNOSIS OF LUNG CANCER

5 REFERENCE TO RELATED APPLICATIONS

This application is related to U.S. Patent Application No. 09/640,878, filed August 18, 2000; U.S. Patent Application No. 09/588,937, filed May 26, 2000; U.S. Patent Application No. 09/538,037, filed March 29, 2000; U.S. Patent Application No. 09/518,809, filed March 3, 2000; U.S. Patent Application No. 09/476,235 filed December 10 30, 1999; U.S. Patent Application No. 09/370,838, filed August 9, 1999; and U.S. Patent Application No. 09/285,323, filed April 2, 1999, each a CIP of the previous application and all pending, and PCT/US00/08560, filed March 30, 2000, pending.

TECHNICAL FIELD OF THE INVENTION

The present invention relates generally to therapy and diagnosis of cancer, 15 such as lung cancer. The invention is more specifically related to polypeptides comprising at least a portion of a lung tumor protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides may be used in compositions for prevention and treatment of lung cancer, and for the diagnosis and monitoring of such cancers.

20 BACKGROUND OF THE INVENTION

Lung cancer is the primary cause of cancer death among both men and women in the U.S., with an estimated 172,000 new cases being reported in 1994. The five-year survival rate among all lung cancer patients, regardless of the stage of disease at diagnosis, is only 13%. This contrasts with a five-year survival rate of 46% among cases 25 detected while the disease is still localized. However, only 16% of lung cancers are discovered before the disease has spread.

Early detection is difficult since clinical symptoms are often not seen until the disease has reached an advanced stage. Currently, diagnosis is aided by the use of chest x-rays, analysis of the type of cells contained in sputum and fiberoptic examination of the bronchial passages. Treatment regimens are determined by the type and stage of the cancer, and include surgery, radiation therapy and/or chemotherapy. In spite of considerable research into therapies for the disease, lung cancer remains difficult to treat.

Accordingly, there remains a need in the art for improved vaccines, treatment methods and diagnostic techniques for lung cancer.

SUMMARY OF THE INVENTION

Briefly stated, the present invention provides compositions and methods for the diagnosis and therapy of cancer, such as lung cancer. In one aspect, the present invention provides polypeptides comprising at least a portion of a lung tumor protein, or a variant thereof. Certain portions and other variants are immunogenic, such that the ability of the variant to react with antigen-specific antisera is not substantially diminished. Within certain embodiments, the polypeptide comprises a sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (a) sequences recited in SEQ ID NOs:217-390, 392, 394, 396, 398-420 and 422-424; (b) variants of a sequence recited in SEQ ID NOs:217-390, 392, 394, 396, 398-420 422-424, 428-433 and 440; and (c) complements of a sequence of (a) or (b). In specific embodiments, the polypeptides of the present invention comprise at least a portion of a tumor protein that includes an amino acid sequence selected from the group consisting of sequences recited in SEQ ID NOs:391, 393, 395, 397, 421, 425-427, 434-439 and variants thereof.

The present invention further provides polynucleotides that encode a polypeptide as described above, or a portion thereof (such as a portion encoding at least 15 amino acid residues of a lung tumor protein), expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, immunogenic
5 compositions, or vaccines for prophylactic or therapeutic use are provided. Such compositions comprise a polypeptide or polynucleotide as described above and an immunostimulant.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a
10 lung tumor protein; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

15 Within related aspects, immunogenic compositions, or vaccines, are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding
20 such fusion proteins.

Within related aspects, pharmaceutical compositions comprising a fusion protein, or a polynucleotide encoding a fusion protein, in combination with a physiologically acceptable carrier are provided.

Compositions are further provided, within other aspects, that comprise a
25 fusion protein, or a polynucleotide encoding a fusion protein, in combination with an immunostimulant.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a composition as recited above. The patient may be afflicted with lung cancer, in which case

the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a lung tumor protein, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a lung tumor protein, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of a lung tumor protein; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for determining the presence or absence of a cancer in a patient, comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody. The cancer may be lung cancer.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a lung tumor protein; (b) detecting in the sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an

oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a lung tumor protein; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

SEQUENCE IDENTIFIERS

SEQ ID NO: 1 is the determined cDNA sequence for L363C1.cons
 SEQ ID NO: 2 is the determined cDNA sequence for L263C2.cons
 SEQ ID NO: 3 is the determined cDNA sequence for L263C2c
 SEQ ID NO: 4 is the determined cDNA sequence for L263C1.cons
 SEQ ID NO: 5 is the determined cDNA sequence for L263C1b
 SEQ ID NO: 6 is the determined cDNA sequence for L164C2.cons
 SEQ ID NO: 7 is the determined cDNA sequence for L164C1.cons
 SEQ ID NO: 8 is the determined cDNA sequence for L366C1a
 SEQ ID NO: 9 is the determined cDNA sequence for L260C1.cons
 SEQ ID NO: 10 is the determined cDNA sequence for L163C1c
 SEQ ID NO: 11 is the determined cDNA sequence for L163C1b

- SEQ ID NO: 12 is the determined cDNA sequence for L255C1.cons
 SEQ ID NO: 13 is the determined cDNA sequence for L255C1b
 SEQ ID NO: 14 is the determined cDNA sequence for L355C1.cons
 SEQ ID NO: 15 is the determined cDNA sequence for L366C1.cons
- 5 SEQ ID NO: 16 is the determined cDNA sequence for L163C1a
 SEQ ID NO: 17 is the determined cDNA sequence for LT86-1
 SEQ ID NO: 18 is the determined cDNA sequence for LT86-2
 SEQ ID NO: 19 is the determined cDNA sequence for LT86-3
 SEQ ID NO: 20 is the determined cDNA sequence for LT86-4
- 10 SEQ ID NO: 21 is the determined cDNA sequence for LT86-5
 SEQ ID NO: 22 is the determined cDNA sequence for LT86-6
 SEQ ID NO: 23 is the determined cDNA sequence for LT86-7
 SEQ ID NO: 24 is the determined cDNA sequence for LT86-8
 SEQ ID NO: 25 is the determined cDNA sequence for LT86-9
- 15 SEQ ID NO: 26 is the determined cDNA sequence for LT86-10
 SEQ ID NO: 27 is the determined cDNA sequence for LT86-11
 SEQ ID NO: 28 is the determined cDNA sequence for LT86-12
 SEQ ID NO: 29 is the determined cDNA sequence for LT86-13
 SEQ ID NO: 30 is the determined cDNA sequence for LT86-14
- 20 SEQ ID NO: 31 is the determined cDNA sequence for LT86-15
 SEQ ID NO: 32 is the predicted amino acid sequence for LT86-1
 SEQ ID NO: 33 is the predicted amino acid sequence for LT86-2
 SEQ ID NO: 34 is the predicted amino acid sequence for LT86-3
 SEQ ID NO: 35 is the predicted amino acid sequence for LT86-4
- 25 SEQ ID NO: 36 is the predicted amino acid sequence for LT86-5
 SEQ ID NO: 37 is the predicted amino acid sequence for LT86-6
 SEQ ID NO: 38 is the predicted amino acid sequence for LT86-7
 SEQ ID NO: 39 is the predicted amino acid sequence for LT86-8
 SEQ ID NO: 40 is the predicted amino acid sequence for LT86-9
- 30 SEQ ID NO: 41 is the predicted amino acid sequence for LT86-10

- SEQ ID NO: 42 is the predicted amino acid sequence for LT86-11
 SEQ ID NO: 43 is the predicted amino acid sequence for LT86-12
 SEQ ID NO: 44 is the predicted amino acid sequence for LT86-13
 SEQ ID NO: 45 is the predicted amino acid sequence for LT86-14
 5 SEQ ID NO: 46 is the predicted amino acid sequence for LT86-15
 SEQ ID NO: 47 is a (dT)₁₂AG primer
 SEQ ID NO: 48 is a primer
 SEQ ID NO: 49 is the determined 5' cDNA sequence for L86S-3
 SEQ ID NO: 50 is the determined 5' cDNA sequence for L86S-12
 10 SEQ ID NO: 51 is the determined 5' cDNA sequence for L86S-16
 SEQ ID NO: 52 is the determined 5' cDNA sequence for L86S-25
 SEQ ID NO: 53 is the determined 5' cDNA sequence for L86S-36
 SEQ ID NO: 54 is the determined 5' cDNA sequence for L86S-40
 SEQ ID NO: 55 is the determined 5' cDNA sequence for L86S-46
 15 SEQ ID NO: 56 is the predicted amino acid sequence for L86S-3
 SEQ ID NO: 57 is the predicted amino acid sequence for L86S-12
 SEQ ID NO: 58 is the predicted amino acid sequence for L86S-16
 SEQ ID NO: 59 is the predicted amino acid sequence for L86S-25
 SEQ ID NO: 60 is the predicted amino acid sequence for L86S-36
 20 SEQ ID NO: 61 is the predicted amino acid sequence for L86S-40
 SEQ ID NO: 62 is the predicted amino acid sequence for L86S-46
 SEQ ID NO: 63 is the determined 5' cDNA sequence for L86S-30
 SEQ ID NO: 64 is the determined 5' cDNA sequence for L86S-41
 SEQ ID NO: 65 is the predicted amino acid sequence from the 5' end of LT86-9
 25 SEQ ID NO: 66 is the determined extended cDNA sequence for LT86-4
 SEQ ID NO: 67 is the predicted extended amino acid sequence for LT86-4
 SEQ ID NO: 68 is the determined 5' cDNA sequence for LT86-20
 SEQ ID NO: 69 is the determined 3' cDNA sequence for LT86-21
 SEQ ID NO: 70 is the determined 5' cDNA sequence for LT86-22
 30 SEQ ID NO: 71 is the determined 5' cDNA sequence for LT86-26

- SEQ ID NO: 72 is the determined 5' cDNA sequence for LT86-27
- SEQ ID NO: 73 is the predicted amino acid sequence for LT86-20
- SEQ ID NO: 74 is the predicted amino acid sequence for LT86-21
- SEQ ID NO: 75 is the predicted amino acid sequence for LT86-22
- 5 SEQ ID NO: 76 is the predicted amino acid sequence for LT86-26
- SEQ ID NO: 77 is the predicted amino acid sequence for LT86-27
- SEQ ID NO: 78 is the determined extended cDNA sequence for L86S-12
- SEQ ID NO: 79 is the determined extended cDNA sequence for L86S-36
- SEQ ID NO: 80 is the determined extended cDNA sequence for L86S-46
- 10 SEQ ID NO: 81 is the predicted extended amino acid sequence for L86S-12
- SEQ ID NO: 82 is the predicted extended amino acid sequence for L86S-36
- SEQ ID NO: 83 is the predicted extended amino acid sequence for L86S-46
- SEQ ID NO: 84 is the determined 5' cDNA sequence for L86S-6
- SEQ ID NO: 85 is the determined 5' cDNA sequence for L86S-11
- 15 SEQ ID NO: 86 is the determined 5' cDNA sequence for L86S-14
- SEQ ID NO: 87 is the determined 5' cDNA sequence for L86S-29
- SEQ ID NO: 88 is the determined 5' cDNA sequence for L86S-34
- SEQ ID NO: 89 is the determined 5' cDNA sequence for L86S-39
- SEQ ID NO: 90 is the determined 5' cDNA sequence for L86S-47
- 20 SEQ ID NO: 91 is the determined 5' cDNA sequence for L86S-49
- SEQ ID NO: 92 is the determined 5' cDNA sequence for L86S-51
- SEQ ID NO: 93 is the predicted amino acid sequence for L86S-6
- SEQ ID NO: 94 is the predicted amino acid sequence for L86S-11
- SEQ ID NO: 95 is the predicted amino acid sequence for L86S-14
- 25 SEQ ID NO: 96 is the predicted amino acid sequence for L86S-29
- SEQ ID NO: 97 is the predicted amino acid sequence for L86S-34
- SEQ ID NO: 98 is the predicted amino acid sequence for L86S-39
- SEQ ID NO: 99 is the predicted amino acid sequence for L86S-47
- SEQ ID NO: 100 is the predicted amino acid sequence for L86S-49
- 30 SEQ ID NO: 101 is the predicted amino acid sequence for L86S-51

30 SEQ ID NO: 131 is the determined cDNA sequence for PSLT-28

- SEQ ID NO: 132 is the determined cDNA sequence for PSLT-30
 SEQ ID NO: 133 is the determined cDNA sequence for PSLT-40
 SEQ ID NO: 134 is the determined cDNA sequence for PSLT-69
 SEQ ID NO: 135 is the determined cDNA sequence for PSLT-71
 5 SEQ ID NO: 136 is the determined cDNA sequence for PSLT-73
 SEQ ID NO: 137 is the determined cDNA sequence for PSLT-79
 SEQ ID NO: 138 is the determined cDNA sequence for PSLT-03
 SEQ ID NO: 139 is the determined cDNA sequence for PSLT-09
 SEQ ID NO: 140 is the determined cDNA sequence for PSLT-011
 10 SEQ ID NO: 141 is the determined cDNA sequence for PSLT-041
 SEQ ID NO: 142 is the determined cDNA sequence for PSLT-62
 SEQ ID NO: 143 is the determined cDNA sequence for PSLT-6
 SEQ ID NO: 144 is the determined cDNA sequence for PSLT-37
 SEQ ID NO: 145 is the determined cDNA sequence for PSLT-74
 15 SEQ ID NO: 146 is the determined cDNA sequence for PSLT-010
 SEQ ID NO: 147 is the determined cDNA sequence for PSLT-012
 SEQ ID NO: 148 is the determined cDNA sequence for PSLT-037
 SEQ ID NO: 149 is the determined 5' cDNA sequence for SAL-3
 SEQ ID NO: 150 is the determined 5' cDNA sequence for SAL-24
 20 SEQ ID NO: 151 is the determined 5' cDNA sequence for SAL-25
 SEQ ID NO: 152 is the determined 5' cDNA sequence for SAL-33
 SEQ ID NO: 153 is the determined 5' cDNA sequence for SAL-50
 SEQ ID NO: 154 is the determined 5' cDNA sequence for SAL-57
 SEQ ID NO: 155 is the determined 5' cDNA sequence for SAL-66
 25 SEQ ID NO: 156 is the determined 5' cDNA sequence for SAL-82
 SEQ ID NO: 157 is the determined 5' cDNA sequence for SAL-99
 SEQ ID NO: 158 is the determined 5' cDNA sequence for SAL-104
 SEQ ID NO: 159 is the determined 5' cDNA sequence for SAL-109
 SEQ ID NO: 160 is the determined 5' cDNA sequence for SAL-5
 30 SEQ ID NO: 161 is the determined 5' cDNA sequence for SAL-8

- SEQ ID NO: 162 is the determined 5' cDNA sequence for SAL-12
 SEQ ID NO: 163 is the determined 5' cDNA sequence for SAL-14
 SEQ ID NO: 164 is the determined 5' cDNA sequence for SAL-16
 SEQ ID NO: 165 is the determined 5' cDNA sequence for SAL-23
 5 SEQ ID NO: 166 is the determined 5' cDNA sequence for SAL-26
 SEQ ID NO: 167 is the determined 5' cDNA sequence for SAL-29
 SEQ ID NO: 168 is the determined 5' cDNA sequence for SAL-32
 SEQ ID NO: 169 is the determined 5' cDNA sequence for SAL-39
 SEQ ID NO: 170 is the determined 5' cDNA sequence for SAL-42
 10 SEQ ID NO: 171 is the determined 5' cDNA sequence for SAL-43
 SEQ ID NO: 172 is the determined 5' cDNA sequence for SAL-44
 SEQ ID NO: 173 is the determined 5' cDNA sequence for SAL-48
 SEQ ID NO: 174 is the determined 5' cDNA sequence for SAL-68
 SEQ ID NO: 175 is the determined 5' cDNA sequence for SAL-72
 15 SEQ ID NO: 176 is the determined 5' cDNA sequence for SAL-77
 SEQ ID NO: 177 is the determined 5' cDNA sequence for SAL-86
 SEQ ID NO: 178 is the determined 5' cDNA sequence for SAL-88
 SEQ ID NO: 179 is the determined 5' cDNA sequence for SAL-93
 SEQ ID NO: 180 is the determined 5' cDNA sequence for SAL-100
 20 SEQ ID NO: 181 is the determined 5' cDNA sequence for SAL-105
 SEQ ID NO: 182 is the predicted amino acid sequence for SAL-3
 SEQ ID NO: 183 is the predicted amino acid sequence for SAL-24
 SEQ ID NO: 184 is a first predicted amino acid sequence for SAL-25
 SEQ ID NO: 185 is a second predicted amino acid sequence for SAL-25
 25 SEQ ID NO: 186 is the predicted amino acid sequence for SAL-33
 SEQ ID NO: 187 is a first predicted amino acid sequence for SAL-50
 SEQ ID NO: 188 is the predicted amino acid sequence for SAL-57
 SEQ ID NO: 189 is a first predicted amino acid sequence for SAL-66
 SEQ ID NO: 190 is a second predicted amino acid sequence for SAL-66
 30 SEQ ID NO: 191 is the predicted amino acid sequence for SAL-82

- SEQ ID NO: 192 is the predicted amino acid sequence for SAL-99
 SEQ ID NO: 193 is the predicted amino acid sequence for SAL-104
 SEQ ID NO: 194 is the predicted amino acid sequence for SAL-5
 SEQ ID NO: 195 is the predicted amino acid sequence for SAL-8
 5 SEQ ID NO: 196 is the predicted amino acid sequence for SAL-12
 SEQ ID NO: 197 is the predicted amino acid sequence for SAL-14
 SEQ ID NO: 198 is the predicted amino acid sequence for SAL-16
 SEQ ID NO: 199 is the predicted amino acid sequence for SAL-23
 SEQ ID NO: 200 is the predicted amino acid sequence for SAL-26
 10 SEQ ID NO: 201 is the predicted amino acid sequence for SAL-29
 SEQ ID NO: 202 is the predicted amino acid sequence for SAL-32
 SEQ ID NO: 203 is the predicted amino acid sequence for SAL-39
 SEQ ID NO: 204 is the predicted amino acid sequence for SAL-42
 SEQ ID NO: 205 is the predicted amino acid sequence for SAL-43
 15 SEQ ID NO: 206 is the predicted amino acid sequence for SAL-44
 SEQ ID NO: 207 is the predicted amino acid sequence for SAL-48
 SEQ ID NO: 208 is the predicted amino acid sequence for SAL-68
 SEQ ID NO: 209 is the predicted amino acid sequence for SAL-72
 SEQ ID NO: 210 is the predicted amino acid sequence for SAL-77
 20 SEQ ID NO: 211 is the predicted amino acid sequence for SAL-86
 SEQ ID NO: 212 is the predicted amino acid sequence for SAL-88
 SEQ ID NO: 213 is the predicted amino acid sequence for SAL-93
 SEQ ID NO: 214 is the predicted amino acid sequence for SAL-100
 SEQ ID NO: 215 is the predicted amino acid sequence for SAL-105
 25 SEQ ID NO: 216 is a second predicted amino acid sequence for SAL-50
 SEQ ID NO: 217 is the determined cDNA sequence for SSLT-4
 SEQ ID NO: 218 is the determined cDNA sequence for SSLT-9
 SEQ ID NO: 219 is the determined cDNA sequence for SSLT-10
 SEQ ID NO: 220 is the determined cDNA sequence for SSLT-12
 30 SEQ ID NO: 221 is the determined cDNA sequence for SSLT-19

- SEQ ID NO: 222 is the determined cDNA sequence for SSLT-31
 SEQ ID NO: 223 is the determined cDNA sequence for SSLT-38
 SEQ ID NO: 224 is the determined cDNA sequence for LT4690-2
 SEQ ID NO: 225 is the determined cDNA sequence for LT4690-3
 5 SEQ ID NO: 226 is the determined cDNA sequence for LT4690-22
 SEQ ID NO: 227 is the determined cDNA sequence for LT4690-24
 SEQ ID NO: 228 is the determined cDNA sequence for LT4690-37
 SEQ ID NO: 229 is the determined cDNA sequence for LT4690-39
 SEQ ID NO: 230 is the determined cDNA sequence for LT4690-40
 10 SEQ ID NO: 231 is the determined cDNA sequence for LT4690-41
 SEQ ID NO: 232 is the determined cDNA sequence for LT4690-49
 SEQ ID NO: 233 is the determined 3' cDNA sequence for LT4690-55
 SEQ ID NO: 234 is the determined 5' cDNA sequence for LT4690-55
 SEQ ID NO: 235 is the determined cDNA sequence for LT4690-59
 15 SEQ ID NO: 236 is the determined cDNA sequence for LT4690-63
 SEQ ID NO: 237 is the determined cDNA sequence for LT4690-71
 SEQ ID NO: 238 is the determined cDNA sequence for 2LT-3
 SEQ ID NO: 239 is the determined cDNA sequence for 2LT-6
 SEQ ID NO: 240 is the determined cDNA sequence for 2LT-22
 20 SEQ ID NO: 241 is the determined cDNA sequence for 2LT-25
 SEQ ID NO: 242 is the determined cDNA sequence for 2LT-26
 SEQ ID NO: 243 is the determined cDNA sequence for 2LT-31
 SEQ ID NO: 244 is the determined cDNA sequence for 2LT-36
 SEQ ID NO: 245 is the determined cDNA sequence for 2LT-42
 25 SEQ ID NO: 246 is the determined cDNA sequence for 2LT-44
 SEQ ID NO: 247 is the determined cDNA sequence for 2LT-54
 SEQ ID NO: 248 is the determined cDNA sequence for 2LT-55
 SEQ ID NO: 249 is the determined cDNA sequence for 2LT-57
 SEQ ID NO: 250 is the determined cDNA sequence for 2LT-58
 30 SEQ ID NO: 251 is the determined cDNA sequence for 2LT-59

SEQ ID NO: 252 is the determined cDNA sequence for 2LT-62
SEQ ID NO: 253 is the determined cDNA sequence for 2LT-63
SEQ ID NO: 254 is the determined cDNA sequence for 2LT-65
SEQ ID NO: 255 is the determined cDNA sequence for 2LT-66
5 SEQ ID NO: 256 is the determined cDNA sequence for 2LT-70
SEQ ID NO: 257 is the determined cDNA sequence for 2LT-73
SEQ ID NO: 258 is the determined cDNA sequence for 2LT-74
SEQ ID NO: 259 is the determined cDNA sequence for 2LT-76
SEQ ID NO: 260 is the determined cDNA sequence for 2LT-77
10 SEQ ID NO: 261 is the determined cDNA sequence for 2LT-78
SEQ ID NO: 262 is the determined cDNA sequence for 2LT-80
SEQ ID NO: 263 is the determined cDNA sequence for 2LT-85
SEQ ID NO: 264 is the determined cDNA sequence for 2LT-87
SEQ ID NO: 265 is the determined cDNA sequence for 2LT-89
15 SEQ ID NO: 266 is the determined cDNA sequence for 2LT-94
SEQ ID NO: 267 is the determined cDNA sequence for 2LT-95
SEQ ID NO: 268 is the determined cDNA sequence for 2LT-98
SEQ ID NO: 269 is the determined cDNA sequence for 2LT-100
SEQ ID NO: 270 is the determined cDNA sequence for 2LT-103
20 SEQ ID NO: 271 is the determined cDNA sequence for 2LT-105
SEQ ID NO: 272 is the determined cDNA sequence for 2LT-107
SEQ ID NO: 273 is the determined cDNA sequence for 2LT-108
SEQ ID NO: 274 is the determined cDNA sequence for 2LT-109
SEQ ID NO: 275 is the determined cDNA sequence for 2LT-118
25 SEQ ID NO: 276 is the determined cDNA sequence for 2LT-120
SEQ ID NO: 277 is the determined cDNA sequence for 2LT-121
SEQ ID NO: 278 is the determined cDNA sequence for 2LT-122
SEQ ID NO: 279 is the determined cDNA sequence for 2LT-124
SEQ ID NO: 280 is the determined cDNA sequence for 2LT-126
30 SEQ ID NO: 281 is the determined cDNA sequence for 2LT-127

- SEQ ID NO: 282 is the determined cDNA sequence for 2LT-128
 SEQ ID NO: 283 is the determined cDNA sequence for 2LT-129
 SEQ ID NO: 284 is the determined cDNA sequence for 2LT-133
 SEQ ID NO: 285 is the determined cDNA sequence for 2LT-137
 5 SEQ ID NO: 286 is the determined cDNA sequence for LT4690-71
 SEQ ID NO: 287 is the determined cDNA sequence for LT4690-82
 SEQ ID NO: 288 is the determined full-length cDNA sequence for SSLT-74
 SEQ ID NO: 289 is the determined cDNA sequence for SSLT-78
 SEQ ID NO: 290 is the determined cDNA sequence for SCC1-8.
 10 SEQ ID NO: 291 is the determined cDNA sequence for SCC1-12.
 SEQ ID NO: 292 is the determined cDNA sequence for SCC1-336
 SEQ ID NO: 293 is the determined cDNA sequence for SCC1-344
 SEQ ID NO: 294 is the determined cDNA sequence for SCC1-345
 SEQ ID NO: 295 is the determined cDNA sequence for SCC1-346
 15 SEQ ID NO: 296 is the determined cDNA sequence for SCC1-348
 SEQ ID NO: 297 is the determined cDNA sequence for SCC1-350
 SEQ ID NO: 298 is the determined cDNA sequence for SCC1-352
 SEQ ID NO: 299 is the determined cDNA sequence for SCC1-354
 SEQ ID NO: 300 is the determined cDNA sequence for SCC1-355
 20 SEQ ID NO: 301 is the determined cDNA sequence for SCC1-356
 SEQ ID NO: 302 is the determined cDNA sequence for SCC1-357
 SEQ ID NO: 303 is the determined cDNA sequence for SCC1-501
 SEQ ID NO: 304 is the determined cDNA sequence for SCC1-503
 SEQ ID NO: 305 is the determined cDNA sequence for SCC1-513
 25 SEQ ID NO: 306 is the determined cDNA sequence for SCC1-516
 SEQ ID NO: 307 is the determined cDNA sequence for SCC1-518
 SEQ ID NO: 308 is the determined cDNA sequence for SCC1-519
 SEQ ID NO: 309 is the determined cDNA sequence for SCC1-522
 SEQ ID NO: 310 is the determined cDNA sequence for SCC1-523
 30 SEQ ID NO: 311 is the determined cDNA sequence for SCC1-525

- SEQ ID NO: 312 is the determined cDNA sequence for SCC1-527
 SEQ ID NO: 313 is the determined cDNA sequence for SCC1-529
 SEQ ID NO: 314 is the determined cDNA sequence for SCC1-530
 SEQ ID NO: 315 is the determined cDNA sequence for SCC1-531
 5 SEQ ID NO: 316 is the determined cDNA sequence for SCC1-532
 SEQ ID NO: 317 is the determined cDNA sequence for SCC1-533
 SEQ ID NO: 318 is the determined cDNA sequence for SCC1-536
 SEQ ID NO: 319 is the determined cDNA sequence for SCC1-538
 SEQ ID NO: 320 is the determined cDNA sequence for SCC1-539
 10 SEQ ID NO: 321 is the determined cDNA sequence for SCC1-541
 SEQ ID NO: 322 is the determined cDNA sequence for SCC1-542
 SEQ ID NO: 323 is the determined cDNA sequence for SCC1-546
 SEQ ID NO: 324 is the determined cDNA sequence for SCC1-549
 SEQ ID NO: 325 is the determined cDNA sequence for SCC1-551
 15 SEQ ID NO: 326 is the determined cDNA sequence for SCC1-552
 SEQ ID NO: 327 is the determined cDNA sequence for SCC1-554
 SEQ ID NO: 328 is the determined cDNA sequence for SCC1-558
 SEQ ID NO: 329 is the determined cDNA sequence for SCC1-559
 SEQ ID NO: 330 is the determined cDNA sequence for SCC1-561
 20 SEQ ID NO: 331 is the determined cDNA sequence for SCC1-562
 SEQ ID NO: 332 is the determined cDNA sequence for SCC1-564
 SEQ ID NO: 333 is the determined cDNA sequence for SCC1-565
 SEQ ID NO: 334 is the determined cDNA sequence for SCC1-566
 SEQ ID NO: 335 is the determined cDNA sequence for SCC1-567
 25 SEQ ID NO: 336 is the determined cDNA sequence for SCC1-568
 SEQ ID NO: 337 is the determined cDNA sequence for SCC1-570
 SEQ ID NO: 338 is the determined cDNA sequence for SCC1-572
 SEQ ID NO: 339 is the determined cDNA sequence for SCC1-575
 SEQ ID NO: 340 is the determined cDNA sequence for SCC1-576
 30 SEQ ID NO: 341 is the determined cDNA sequence for SCC1-577

- SEQ ID NO: 342 is the determined cDNA sequence for SCC1-578
 SEQ ID NO: 343 is the determined cDNA sequence for SCC1-582
 SEQ ID NO: 344 is the determined cDNA sequence for SCC1-583
 SEQ ID NO: 345 is the determined cDNA sequence for SCC1-586
 5 SEQ ID NO: 346 is the determined cDNA sequence for SCC1-588
 SEQ ID NO: 347 is the determined cDNA sequence for SCC1-590
 SEQ ID NO: 348 is the determined cDNA sequence for SCC1-591
 SEQ ID NO: 349 is the determined cDNA sequence for SCC1-592
 SEQ ID NO: 350 is the determined cDNA sequence for SCC1-593
 10 SEQ ID NO: 351 is the determined cDNA sequence for SCC1-594
 SEQ ID NO: 352 is the determined cDNA sequence for SCC1-595
 SEQ ID NO: 353 is the determined cDNA sequence for SCC1-596
 SEQ ID NO: 354 is the determined cDNA sequence for SCC1-598
 SEQ ID NO: 355 is the determined cDNA sequence for SCC1-599
 15 SEQ ID NO: 356 is the determined cDNA sequence for SCC1-602
 SEQ ID NO: 357 is the determined cDNA sequence for SCC1-604
 SEQ ID NO: 358 is the determined cDNA sequence for SCC1-605
 SEQ ID NO: 359 is the determined cDNA sequence for SCC1-606
 SEQ ID NO: 360 is the determined cDNA sequence for SCC1-607
 20 SEQ ID NO: 361 is the determined cDNA sequence for SCC1-608
 SEQ ID NO: 362 is the determined cDNA sequence for SCC1-610
 SEQ ID NO: 363 is the determined cDNA sequence for clone DMS79T1
 SEQ ID NO: 364 is the determined cDNA sequence for clone DMS79T2
 SEQ ID NO: 365 is the determined cDNA sequence for clone DMS79T3
 25 SEQ ID NO: 366 is the determined cDNA sequence for clone DMS79T5
 SEQ ID NO: 367 is the determined cDNA sequence for clone DMS79T6
 SEQ ID NO: 368 is the determined cDNA sequence for clone DMS79T7
 SEQ ID NO: 369 is the determined cDNA sequence for clone DMS79T9
 SEQ ID NO: 370 is the determined cDNA sequence for clone DMS79T10
 30 SEQ ID NO: 371 is the determined cDNA sequence for clone DMS79T11

- SEQ ID NO: 372 is the determined cDNA sequence for clone 128T1
 SEQ ID NO: 373 is the determined cDNA sequence for clone 128T2
 SEQ ID NO: 374 is the determined cDNA sequence for clone 128T3
 SEQ ID NO: 375 is the determined cDNA sequence for clone 128T4
 5 SEQ ID NO: 376 is the determined cDNA sequence for clone 128T5
 SEQ ID NO: 377 is the determined cDNA sequence for clone 128T7
 SEQ ID NO: 378 is the determined cDNA sequence for clone 128T9
 SEQ ID NO: 379 is the determined cDNA sequence for clone 128T10
 SEQ ID NO: 380 is the determined cDNA sequence for clone 128T11
 10 SEQ ID NO: 381 is the determined cDNA sequence for clone 128T12
 SEQ ID NO: 382 is the determined cDNA sequence for clone NCIH69T3
 SEQ ID NO: 383 is the determined cDNA sequence for clone NCIH69T5
 SEQ ID NO: 384 is the determined cDNA sequence for clone NCIH69T6
 SEQ ID NO: 385 is the determined cDNA sequence for clone NCIH69T7
 15 SEQ ID NO: 386 is the determined cDNA sequence for clone NCIH69T9
 SEQ ID NO: 387 is the determined cDNA sequence for clone NCIH69T10
 SEQ ID NO: 388 is the determined cDNA sequence for clone NCIH69T11
 SEQ ID NO: 389 is the determined cDNA sequence for clone NCIH69T12
 SEQ ID NO: 390 is the full-length cDNA sequence for 128T1
 20 SEQ ID NO: 391 is the amino acid sequence for 128T1
 SEQ ID NO: 392 is the full-length cDNA sequence for 2LT-128
 SEQ ID NO: 393 is the amino acid sequence for 2LT-128
 SEQ ID NO: 394 is an extended cDNA sequence for clone SCC1-542
 SEQ ID NO: 395 is the amino acid sequence corresponding to SEQ ID NO:394
 25 SEQ ID NO: 396 is an extended cDNA sequence for clone SCC1-593
 SEQ ID NO: 397 is the amino acid sequence corresponding to SEQ ID NO:396
 SEQ ID NO:398 is the determined cDNA sequence for 55508.1
 SEQ ID NO:399 is the determined cDNA sequence for 55509.1
 SEQ ID NO:400 is the determined cDNA sequence for 54243.1
 30 SEQ ID NO:401 is the determined cDNA sequence for 54251.1

- SEQ ID NO:402 is the determined cDNA sequence for 54252.1
 SEQ ID NO:403 is the determined cDNA sequence for 54253.1
 SEQ ID NO:404 is the determined cDNA sequence for 55518.1
 SEQ ID NO:405 is the determined cDNA sequence for 54258.1
 5 SEQ ID NO:406 is the determined cDNA sequence for 54575.1
 SEQ ID NO:407 is the determined cDNA sequence for 54577.1
 SEQ ID NO:408 is the determined cDNA sequence for 54584.1
 SEQ ID NO:409 is the determined cDNA sequence for 55521.1
 SEQ ID NO:410 is the determined cDNA sequence for 54589.1
 10 SEQ ID NO:411 is the determined cDNA sequence for 54592.1
 SEQ ID NO:412 is the determined cDNA sequence for 55134.1
 SEQ ID NO:413 is the determined cDNA sequence for 55137.1
 SEQ ID NO:414 is the determined cDNA sequence for 55140.1
 SEQ ID NO:415 is the determined cDNA sequence for 55531.1
 15 SEQ ID NO:416 is the determined cDNA sequence for 55532.1
 SEQ ID NO:417 is the determined cDNA sequence for 54621.1
 SEQ ID NO:418 is the determined cDNA sequence for 55548.1
 SEQ ID NO:419 is the determined cDNA sequence for 54623.1
 SEQ ID NO:420 is the determined cDNA sequence for L39
 20 SEQ ID NO:421 is the predicted amino acid sequence for L39
 SEQ ID NO:422 is the determined cDNA sequence for SCC2-29
 SEQ ID NO:423 is the determined cDNA sequence for SCC2-36
 SEQ ID NO:424 is the determined cDNA sequence for SCC2-60
 SEQ ID NO:425 is the predicted amino acid sequence for SCC2-29
 25 SEQ ID NO:426 is the predicted amino acid sequence for SCC2-36
 SEQ ID NO:427 is the predicted amino acid sequence for SCC2-60
 SEQ ID NO:428 is an extended cDNA sequence for the clone 20129, also referred to as
 2LT-3, set forth in SEQ ID NO: 238
 SEQ ID NO:429 is an extended cDNA sequence for the clone 20347, also referred to as
 30 2LT-26, set forth in SEQ ID NO: 242

SEQ ID NO:430 is an extended cDNA sequence for the clone 21282, also referred to as
2LT-57, set forth in SEQ ID NO: 249

SEQ ID NO:431 is an extended cDNA sequence for the clone 21283, also referred to as
2LT-58, set forth in SEQ ID NO: 250

5 SEQ ID NO:432 is an extended cDNA sequence for the clone 21484, also referred to as
2LT-98, set forth in SEQ ID NO: 268

SEQ ID NO:433 is an extended cDNA sequence for the clone 21871, also referred to as
2LT-124, set forth in SEQ ID NO: 279

SEQ ID NO:434 is an amino acid sequence encoded by SEQ ID NO: 428

10 SEQ ID NO:435 is an amino acid sequence encoded by SEQ ID NO: 429

SEQ ID NO:436 is an amino acid sequence encoded by SEQ ID NO: 430

SEQ ID NO:437 is an amino acid sequence encoded by SEQ ID NO: 431

SEQ ID NO:438 is an amino acid sequence encoded by SEQ ID NO: 432

SEQ ID NO:439 is an amino acid sequence encoded by SEQ ID NO: 433

15 SEQ ID NO:440 is the determined cDNA sequence for clone 19A4

DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to compositions
20 and methods for using the compositions, for example in the therapy and diagnosis of
cancer, such as lung cancer. Certain illustrative compositions described herein include
lung tumor polypeptides, polynucleotides encoding such polypeptides, binding agents such
as antibodies, antigen presenting cells (APCs) and/or immune system cells (*e.g.*, T cells).
A "lung tumor protein," as the term is used herein, refers generally to a protein that is
25 expressed in lung tumor cells at a level that is at least two fold, and preferably at least five
fold, greater than the level of expression in a normal tissue, as determined using a
representative assay provided herein. Certain lung tumor proteins are tumor proteins that

react detectably (within an immunoassay, such as an ELISA or Western blot) with antisera of a patient afflicted with lung cancer.

Therefore, in accordance with the above, and as described further below, the present invention provides illustrative polynucleotide compositions having sequences set forth in SEQ ID NOs:217-390, 392, 394, 396, 398-420, 422-424, 428-433 and 440 illustrative polypeptide compositions having amino acid sequences set forth in SEQ ID NOs:391, 393, 395 and 397, 421, 425-427 and 434-439, antibody compositions capable of binding such polypeptides, and numerous additional embodiments employing such compositions, for example in the detection, diagnosis and/or therapy of human lung cancer.

10 POLYNUCLEOTIDE COMPOSITIONS

As used herein, the terms "DNA segment" and "polynucleotide" refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. Therefore, a DNA segment encoding a polypeptide refers to a DNA segment that contains one or more coding sequences yet is substantially isolated away from, or purified free from, total genomic DNA of the species from which the DNA segment is obtained. Included within the terms "DNA segment" and "polynucleotide" are DNA segments and smaller fragments of such segments, and also recombinant vectors, including, for example, plasmids, cosmids, phagemids, phage, viruses, and the like.

As will be understood by those skilled in the art, the DNA segments of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

"Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA segment does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA segment as originally

isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be recognized by the skilled artisan, polynucleotides may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a lung tumor protein or a portion thereof) or may comprise a variant, or a biological or antigenic functional equivalent of such a sequence. Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions, as further described below, preferably such that the immunogenicity of the encoded polypeptide is not diminished, relative to a native tumor protein. The effect on the immunogenicity of the encoded polypeptide may generally be assessed as described herein. The term “variants” also encompasses homologous genes of xenogenic origin.

When comparing polynucleotide or polypeptide sequences, two sequences are said to be “identical” if the sequence of nucleotides or amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A “comparison window” as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment

schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified
 5 Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of*
 10 *Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA* 80:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.*
 15 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining
 20 percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.* (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software for
 25 performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix can be used to calculate the cumulative score.

Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Therefore, the present invention encompasses polynucleotide and polypeptide sequences having substantial identity to the sequences disclosed herein, for example those comprising at least 50% sequence identity, preferably at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence identity compared to a polynucleotide or polypeptide sequence of this invention using the methods described herein, (*e.g.*, BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide

sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

In additional embodiments, the present invention provides isolated polynucleotides and polypeptides comprising various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at least about 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative DNA segments with total lengths of about 10,000, about 5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like, (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

In other embodiments, the present invention is directed to polynucleotides that are capable of hybridizing under moderately stringent conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a

solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS.

Moreover, it will be appreciated by those of ordinary skill in the art that, as
 5 a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided
 10 herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

15 **PROBES AND PRIMERS**

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a sequence region of at least about 15 nucleotide long contiguous sequence that has the same
 20 sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

25 The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned,

such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequence set forth in SEQ ID NOs:217-390, 392, 394, 396, 398-420, 422-424, 428-433 and 440, or to any continuous portion of the sequence, from about 15-25 nucleotides in length up to and including the full length sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR™ technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be

readily manipulated, and thus will generally be a method of choice depending on the desired results.

POLYNUCLEOTIDE IDENTIFICATION AND CHARACTERIZATION

Polynucleotides may be identified, prepared and/or manipulated using any
 5 of a variety of well established techniques. For example, a polynucleotide may be
 identified, as described in more detail below, by screening a microarray of cDNAs for
 tumor-associated expression (*i.e.*, expression that is at least two fold greater in a tumor than
 in normal tissue, as determined using a representative assay provided herein). Such screens
 may be performed, for example, using a Synteni microarray (Palo Alto, CA) according to
 10 the manufacturer's instructions (and essentially as described by Schena *et al.*, *Proc. Natl.*
Acad. Sci. USA 93:10614-10619, 1996 and Heller *et al.*, *Proc. Natl. Acad. Sci. USA*
 94:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA
 prepared from cells expressing the proteins described herein, such as lung tumor cells.
 Such polynucleotides may be amplified via polymerase chain reaction (PCR). For this
 15 approach, sequence-specific primers may be designed based on the sequences provided
 herein, and may be purchased or synthesized.

An amplified portion of a polynucleotide of the present invention may be
 used to isolate a full length gene from a suitable library (*e.g.*, a lung tumor cDNA library)
 using well known techniques. Within such techniques, a library (cDNA or genomic) is
 20 screened using one or more polynucleotide probes or primers suitable for amplification.
 Preferably, a library is size-selected to include larger molecules. Random primed libraries
 may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries
 are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (*e.g.*, by
 25 nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or
 bacteriophage library is then generally screened by hybridizing filters containing denatured
 bacterial colonies (or lawns containing phage plaques) with the labeled probe (*see*
 Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor

Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can then be assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

Alternatively, there are numerous amplification techniques for obtaining a full length coding sequence from a partial cDNA sequence. Within such techniques, amplification is generally performed via PCR. Any of a variety of commercially available kits may be used to perform the amplification step. Primers may be designed using, for example, software well known in the art. Primers are preferably 22-30 nucleotides in length, have a GC content of at least 50% and anneal to the target sequence at temperatures of about 68°C to 72°C. The amplified region may be sequenced as described above, and overlapping sequences assembled into a contiguous sequence.

One such amplification technique is inverse PCR (*see* Triglia *et al.*, *Nucl. Acids Res.* 16:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector

sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom *et al.*, *PCR Methods Applic. 1*:111-19, 1991) and walking PCR (Parker *et al.*, *Nucl. Acids. Res. 19*:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

5 In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (*e.g.*, NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences may also be
10 obtained by analysis of genomic fragments.

POLYNUCLEOTIDE EXPRESSION IN HOST CELLS

 In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct
15 expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

 As will be understood by those of skill in the art, it may be advantageous in
20 some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

25 Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA

shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. *et al.* (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. *et al.* (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. *et al.* (1995) *Science* 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, T. (1983) *Proteins, Structures and Molecular Principles*, WH Freeman and Co., New York, N.Y.) or other comparable techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. *et al.* (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. *et al.* (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York. N.Y.

A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (*e.g.*, baculovirus); plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are generally preferred. If it is necessary to

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β -galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. *et al.* (1984) *EMBO J.* 3:1671-1680;

An insect system may also be used to express a polypeptide of interest. For example, in one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. *et al.* (1994) *Proc. Natl. Acad. Sci.* 91 :3224-3227).

In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate

expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. *et al.* (1994) *Results Probl. Cell Differ.* 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. *et al.* (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. *et al.* (1990) *Cell* 22:817-23) genes which can be employed in tk.sup.- or aprt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. *et al.* (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. *et al.* (1981) *J. Mol. Biol.* 150:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. *et al.* (1995) *Methods Mol. Biol.* 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells which contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations

and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies
 5 specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and
 10 other assays are described, among other places, in Hampton, R. *et al.* (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, Minn.) and Maddox, D. E. *et al.* (1983; *J. Exp. Med.* 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for
 15 producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by
 20 addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

25 Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may

be designed to contain signal sequences which direct secretion of the encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such

5 purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or

10 enterokinase (Invitrogen. San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal

15 ion affinity chromatography) as described in Porath, J. *et al.* (1992, *Prot. Exp. Purif.* 3:263-281) while the enterokinase cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. *et al.* (1993; *DNA Cell Biol.* 12:441-453).

In addition to recombinant production methods, polypeptides of the

20 invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and

25 combined using chemical methods to produce the full length molecule.

SITE-SPECIFIC MUTAGENESIS

Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent polypeptides, through specific

mutagenesis of the underlying polynucleotides that encode them. The technique, well-known to those of skill in the art, further provides a ready ability to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific
5 mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve,
10 alter, decrease, modify, or otherwise change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties
15 of the encoded polypeptide, such as the antigenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in
20 length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors
25 such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is
5 prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to
10 transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants
15 of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein
20 by reference, for that purpose.

As used herein, the term “oligonucleotide directed mutagenesis procedure” refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as
25 amplification. As used herein, the term “oligonucleotide directed mutagenesis procedure” is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example,

Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in
 5 its entirety.

POLYNUCLEOTIDE AMPLIFICATION TECHNIQUES

A number of template dependent processes are available to amplify the target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCR™) which is described in detail in U.S.
 10 Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCR™, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (*e.g.*, *Taq* polymerase). If the target sequence is present in a sample, the
 15 primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCR™ amplification
 20 procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

Another method for amplification is the ligase chain reaction (referred to as LCR), disclosed in Eur. Pat. Appl. Publ. No. 320,308 (specifically incorporated herein by reference in its entirety). In LCR, two complementary probe pairs are prepared, and in the
 25 presence of the target sequence, each pair will bind to opposite complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCR™, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S.

Patent No. 4,883,750, incorporated herein by reference in its entirety, describes an alternative method of amplification similar to LCR for binding probe pairs to a target sequence.

Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880, incorporated herein by reference in its entirety, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA that has a region complementary to that of a target is added to a sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence that can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[α -thio]triphosphates in one strand of a restriction site (Walker *et al.*, 1992, incorporated herein by reference in its entirety), may also be useful in the amplification of nucleic acids in the present invention.

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, *i.e.* nick translation. A similar method, called Repair Chain Reaction (RCR) is another method of amplification which may be useful in the present invention and is involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA.

Sequences can also be detected using a cyclic probe reaction (CPR). In CPR, a probe having a 3' and 5' sequences of non-target DNA and an internal or "middle" sequence of the target protein specific RNA is hybridized to DNA which is present in a sample. Upon hybridization, the reaction is treated with RNaseH, and the products of the probe are identified as distinctive products by generating a signal that is released after digestion. The original template is annealed to another cycling probe and the reaction is

Still other amplification methods described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025, each of which is
5 incorporated herein by reference in its entirety, may be used in accordance with the present invention. In the former application, "modified" primers are used in a PCR-like, template and enzyme dependent synthesis. The primers may be modified by labeling with a capture moiety (*e.g.*, biotin) and/or a detector moiety (*e.g.*, enzyme). In the latter application, an excess of labeled probes is added to a sample. In the presence of the target sequence, the
10 probe binds and is cleaved catalytically. After cleavage, the target sequence is released intact to be bound by excess probe. Cleavage of the labeled probe signals the presence of the target sequence.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (Kwoh *et al.*, 1989; PCT Intl. Pat. Appl. Publ. No. WO 88/10315, incorporated herein by reference in its entirety), including nucleic acid sequence based amplification (NASBA) and 3SR. In NASBA, the nucleic acids can be prepared for amplification by standard phenol/chloroform extraction, heat denaturation of a sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride extraction of RNA. These amplification techniques involve annealing a primer that has sequences specific to the target sequence. Following polymerization, DNA/RNA hybrids are digested with RNase H while double stranded DNA molecules are heat-denatured again. In either case the single stranded DNA is made fully double stranded by addition of second target-specific primer, followed by polymerization. The double stranded DNA molecules are then multiply transcribed by a polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNAs are reverse transcribed into DNA, and transcribed once again with a polymerase such as T7 or SP6. The resulting products, whether truncated or complete, indicate target-specific sequences.

Eur. Pat. Appl. Publ. No. 329,822, incorporated herein by reference in its entirety, disclose a nucleic acid amplification process involving cyclically synthesizing

single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention. The ssRNA is a first template for a first primer oligonucleotide, which is elongated by reverse transcriptase (RNA-dependent DNA polymerase). The RNA is then removed from resulting DNA:RNA duplex by the action of

5 ribonuclease H (RNase H, an RNase specific for RNA in a duplex with either DNA or RNA). The resultant ssDNA is a second template for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerase) 5' to its homology to its template. This primer is then extended by DNA polymerase (exemplified by the large "Klenow" fragment of *E. coli* DNA polymerase I), resulting as a

10 double-stranded DNA ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having additionally, at one end, a promoter sequence. This promoter sequence can be used by the appropriate RNA polymerase to make many RNA copies of the DNA. These copies can then re-enter the cycle leading to very swift amplification. With proper choice of enzymes, this amplification can be done

15 isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be chosen to be in the form of either DNA or RNA.

PCT Intl. Pat. Appl. Publ. No. WO 89/06700, incorporated herein by reference in its entirety, disclose a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA")

20 followed by transcription of many RNA copies of the sequence. This scheme is not cyclic; *i.e.* new templates are not produced from the resultant RNA transcripts. Other amplification methods include "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) which are well-known to those of skill in the art.

Methods based on ligation of two (or more) oligonucleotides in the presence

25 of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying the di-oligonucleotide (Wu and Dean, 1996, incorporated herein by reference in its entirety), may also be used in the amplification of DNA sequences of the present invention.

BIOLOGICAL FUNCTIONAL EQUIVALENTS

Modification and changes may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a polypeptide with desirable characteristics. As mentioned above, it is often desirable to introduce one or more mutations into a specific polynucleotide sequence. In certain circumstances, the resulting encoded polypeptide sequence is altered by this mutation, or in other cases, the sequence of the polypeptide is unchanged by one or more mutations in the encoding polynucleotide.

When it is desirable to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, second-generation molecule, the amino acid changes may be achieved by changing one or more of the codons of the encoding DNA sequence, according to Table 1.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

TABLE 1

Amino Acids			Codons					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive
5 biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporated herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been

assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (−0.4); threonine (−0.7); serine (−0.8); tryptophan (−0.9); tyrosine (−1.3); proline (−1.6); histidine (−3.2); glutamate (−3.5); glutamine (−3.5); aspartate (−3.5); asparagine (−3.5); lysine (−3.9); and arginine (−4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 \pm 1); glutamate (+3.0 \pm 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (−0.4); proline (−0.5 \pm 1); alanine (−0.5); histidine (−0.5); cysteine (−1.0); methionine (−1.3); valine (−1.5); leucine (−1.8); isoleucine (−1.8); tyrosine (−2.3); phenylalanine (−2.5); tryptophan (−3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their

hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

5 In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl- methyl-, thio- and other
10 modified forms of adenine, cytidine, guanine, thymine and uridine.

IN VIVO POLYNUCLEOTIDE DELIVERY TECHNIQUES

 In additional embodiments, genetic constructs comprising one or more of the polynucleotides of the invention are introduced into cells *in vivo*. This may be achieved using any of a variety of well known approaches, several of which are outlined below for
15 the purpose of illustration.

1. ADENOVIRUS

 One of the preferred methods for *in vivo* delivery of one or more nucleic acid sequences involves the use of an adenovirus expression vector. "Adenovirus expression vector" is meant to include those constructs containing adenovirus sequences
20 sufficient to (a) support packaging of the construct and (b) to express a polynucleotide that has been cloned therein in a sense or antisense orientation. Of course, in the context of an antisense construct, expression does not require that the gene product be synthesized.

 The expression vector comprises a genetically engineered form of an adenovirus. Knowledge of the genetic organization of adenovirus, a 36 kb, linear, double-
25 stranded DNA virus, allows substitution of large pieces of adenoviral DNA with foreign sequences up to 7 kb (Grunhaus and Horwitz, 1992). In contrast to retrovirus, the adenoviral infection of host cells does not result in chromosomal integration because

adenoviral DNA can replicate in an episomal manner without potential genotoxicity. Also, adenoviruses are structurally stable, and no genome rearrangement has been detected after extensive amplification. Adenovirus can infect virtually all epithelial cells regardless of their cell cycle stage. So far, adenoviral infection appears to be linked only to mild disease
 5 such as acute respiratory disease in humans.

Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized genome, ease of manipulation, high titer, wide target-cell range and high infectivity. Both ends of the viral genome contain 100-200 base pair inverted repeats (ITRs), which are *cis* elements necessary for viral DNA replication and packaging. The
 10 early (E) and late (L) regions of the genome contain different transcription units that are divided by the onset of viral DNA replication. The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are involved in DNA replication,
 15 late gene expression and host cell shut-off (Renan, 1990). The products of the late genes, including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter (MLP). The MLP, (located at 16.8 m.u.) is particularly efficient during the late phase of infection, and all the mRNA's issued from this promoter possess a 5'-tripartite leader (TPL) sequence
 20 which makes them preferred mRNA's for translation.

In a current system, recombinant adenovirus is generated from homologous recombination between shuttle vector and provirus vector. Due to the possible recombination between two proviral vectors, wild-type adenovirus may be generated from this process. Therefore, it is critical to isolate a single clone of virus from an individual
 25 plaque and examine its genomic structure.

Generation and propagation of the current adenovirus vectors, which are replication deficient, depend on a unique helper cell line, designated 293, which was transformed from human embryonic kidney cells by Ad5 DNA fragments and constitutively expresses E1 proteins (Graham *et al.*, 1977). Since the E3 region is

dispensable from the adenovirus genome (Jones and Shenk, 1978), the current adenovirus vectors, with the help of 293 cells, carry foreign DNA in either the E1, the D3 or both regions (Graham and Prevec, 1991). In nature, adenovirus can package approximately 105% of the wild-type genome (Ghosh-Choudhury *et al.*, 1987), providing capacity for about 2 extra kB of DNA. Combined with the approximately 5.5 kB of DNA that is replaceable in the E1 and E3 regions, the maximum capacity of the current adenovirus vector is under 7.5 kB, or about 15% of the total length of the vector. More than 80% of the adenovirus viral genome remains in the vector backbone and is the source of vector-borne cytotoxicity. Also, the replication deficiency of the E1-deleted virus is incomplete. For example, leakage of viral gene expression has been observed with the currently available vectors at high multiplicities of infection (MOI) (Mulligan, 1993).

Helper cell lines may be derived from human cells such as human embryonic kidney cells, muscle cells, hematopoietic cells or other human embryonic mesenchymal or epithelial cells. Alternatively, the helper cells may be derived from the cells of other mammalian species that are permissive for human adenovirus. Such cells include, *e.g.*, Vero cells or other monkey embryonic mesenchymal or epithelial cells. As stated above, the currently preferred helper cell line is 293.

Recently, Racher *et al.* (1995) disclosed improved methods for culturing 293 cells and propagating adenovirus. In one format, natural cell aggregates are grown by inoculating individual cells into 1 liter siliconized spinner flasks (Techne, Cambridge, UK) containing 100-200 ml of medium. Following stirring at 40 rpm, the cell viability is estimated with trypan blue. In another format, Fibra-Cel microcarriers (Bibby Sterlin, Stone, UK) (5 g/l) is employed as follows. A cell inoculum, resuspended in 5 ml of medium, is added to the carrier (50 ml) in a 250 ml Erlenmeyer flask and left stationary, with occasional agitation, for 1 to 4 h. The medium is then replaced with 50 ml of fresh medium and shaking initiated. For virus production, cells are allowed to grow to about 80% confluence, after which time the medium is replaced (to 25% of the final volume) and adenovirus added at an MOI of 0.05. Cultures are left stationary overnight, following which the volume is increased to 100% and shaking commenced for another 72 h.

Other than the requirement that the adenovirus vector be replication defective, or at least conditionally defective, the nature of the adenovirus vector is not believed to be crucial to the successful practice of the invention. The adenovirus may be of any of the 42 different known serotypes or subgroups A-F. Adenovirus type 5 of subgroup C is the preferred starting material in order to obtain a conditional replication-defective adenovirus vector for use in the present invention, since Adenovirus type 5 is a human adenovirus about which a great deal of biochemical and genetic information is known, and it has historically been used for most constructions employing adenovirus as a vector.

As stated above, the typical vector according to the present invention is replication defective and will not have an adenovirus E1 region. Thus, it will be most convenient to introduce the polynucleotide encoding the gene of interest at the position from which the E1-coding sequences have been removed. However, the position of insertion of the construct within the adenovirus sequences is not critical to the invention. The polynucleotide encoding the gene of interest may also be inserted in lieu of the deleted E3 region in E3 replacement vectors as described by Karlsson *et al.* (1986) or in the E4 region where a helper cell line or helper virus complements the E4 defect.

Adenovirus is easy to grow and manipulate and exhibits broad host range *in vitro* and *in vivo*. This group of viruses can be obtained in high titers, *e.g.*, 10^9 - 10^{11} plaque-forming units per ml, and they are highly infective. The life cycle of adenovirus does not require integration into the host cell genome. The foreign genes delivered by adenovirus vectors are episomal and, therefore, have low genotoxicity to host cells. No side effects have been reported in studies of vaccination with wild-type adenovirus (Couch *et al.*, 1963; Top *et al.*, 1971), demonstrating their safety and therapeutic potential as *in vivo* gene transfer vectors.

Adenovirus vectors have been used in eukaryotic gene expression (Levrero *et al.*, 1991; Gomez-Foix *et al.*, 1992) and vaccine development (Grunhaus and Horwitz, 1992; Graham and Prevec, 1992). Recently, animal studies suggested that recombinant adenovirus could be used for gene therapy (Stratford-Perricaudet and Perricaudet, 1991; Stratford-Perricaudet *et al.*, 1990; Rich *et al.*, 1993). Studies in administering recombinant

adenovirus to different tissues include trachea instillation (Rosenfeld *et al.*, 1991; Rosenfeld *et al.*, 1992), muscle injection (Ragot *et al.*, 1993), peripheral intravenous injections (Herz and Gerard, 1993) and stereotactic inoculation into the brain (Le Gal La Salle *et al.*, 1993).

5 2. RETROVIRUSES

The retroviruses are a group of single-stranded RNA viruses characterized by an ability to convert their RNA to double-stranded DNA in infected cells by a process of reverse-transcription (Coffin, 1990). The resulting DNA then stably integrates into cellular chromosomes as a provirus and directs synthesis of viral proteins. The integration results in the retention of the viral gene sequences in the recipient cell and its descendants. The retroviral genome contains three genes, gag, pol, and env that code for capsid proteins, polymerase enzyme, and envelope components, respectively. A sequence found upstream from the gag gene contains a signal for packaging of the genome into virions. Two long terminal repeat (LTR) sequences are present at the 5' and 3' ends of the viral genome. These contain strong promoter and enhancer sequences and are also required for integration in the host cell genome (Coffin, 1990).

In order to construct a retroviral vector, a nucleic acid encoding one or more oligonucleotide or polynucleotide sequences of interest is inserted into the viral genome in the place of certain viral sequences to produce a virus that is replication-defective. In order to produce virions, a packaging cell line containing the gag, pol, and env genes but without the LTR and packaging components is constructed (Mann *et al.*, 1983). When a recombinant plasmid containing a cDNA, together with the retroviral LTR and packaging sequences is introduced into this cell line (by calcium phosphate precipitation for example), the packaging sequence allows the RNA transcript of the recombinant plasmid to be packaged into viral particles, which are then secreted into the culture media (Nicolas and Rubenstein, 1988; Temin, 1986; Mann *et al.*, 1983). The media containing the recombinant retroviruses is then collected, optionally concentrated, and used for gene

transfer. Retroviral vectors are able to infect a broad variety of cell types. However, integration and stable expression require the division of host cells (Paskind *et al.*, 1975).

A novel approach designed to allow specific targeting of retrovirus vectors was recently developed based on the chemical modification of a retrovirus by the chemical addition of lactose residues to the viral envelope. This modification could permit the specific infection of hepatocytes *via* sialoglycoprotein receptors.

A different approach to targeting of recombinant retroviruses was designed in which biotinylated antibodies against a retroviral envelope protein and against a specific cell receptor were used. The antibodies were coupled *via* the biotin components by using streptavidin (Roux *et al.*, 1989). Using antibodies against major histocompatibility complex class I and class II antigens, they demonstrated the infection of a variety of human cells that bore those surface antigens with an ecotropic virus *in vitro* (Roux *et al.*, 1989).

3. ADENO-ASSOCIATED VIRUSES

AAV (Ridgeway, 1988; Hermonat and Muzyczka, 1984) is a parovirus, discovered as a contamination of adenoviral stocks. It is a ubiquitous virus (antibodies are present in 85% of the US human population) that has not been linked to any disease. It is also classified as a dependovirus, because its replications is dependent on the presence of a helper virus, such as adenovirus. Five serotypes have been isolated, of which AAV-2 is the best characterized. AAV has a single-stranded linear DNA that is encapsidated into capsid proteins VP1, VP2 and VP3 to form an icosahedral virion of 20 to 24 nm in diameter (Muzyczka and McLaughlin, 1988).

The AAV DNA is approximately 4.7 kilobases long. It contains two open reading frames and is flanked by two ITRs. There are two major genes in the AAV genome: *rep* and *cap*. The *rep* gene codes for proteins responsible for viral replications, whereas *cap* codes for capsid protein VP1-3. Each ITR forms a T-shaped hairpin structure. These terminal repeats are the only essential *cis* components of the AAV for chromosomal integration. Therefore, the AAV can be used as a vector with all viral coding sequences removed and replaced by the cassette of genes for delivery. Three viral

promoters have been identified and named p5, p19, and p40, according to their map position. Transcription from p5 and p19 results in production of rep proteins, and transcription from p40 produces the capsid proteins (Hermonat and Muzyczka, 1984).

There are several factors that prompted researchers to study the possibility of using rAAV as an expression vector. One is that the requirements for delivering a gene to integrate into the host chromosome are surprisingly few. It is necessary to have the 145-bp ITRs, which are only 6% of the AAV genome. This leaves room in the vector to assemble a 4.5-kb DNA insertion. While this carrying capacity may prevent the AAV from delivering large genes, it is amply suited for delivering the antisense constructs of the present invention.

AAV is also a good choice of delivery vehicles due to its safety. There is a relatively complicated rescue mechanism: not only wild type adenovirus but also AAV genes are required to mobilize rAAV. Likewise, AAV is not pathogenic and not associated with any disease. The removal of viral coding sequences minimizes immune reactions to viral gene expression, and therefore, rAAV does not evoke an inflammatory response.

4. OTHER VIRAL VECTORS AS EXPRESSION CONSTRUCTS

Other viral vectors may be employed as expression constructs in the present invention for the delivery of oligonucleotide or polynucleotide sequences to a host cell. Vectors derived from viruses such as vaccinia virus (Ridgeway, 1988; Coupar *et al.*, 1988), lentiviruses, polio viruses and herpes viruses may be employed. They offer several attractive features for various mammalian cells (Friedmann, 1989; Ridgeway, 1988; Coupar *et al.*, 1988; Horwich *et al.*, 1990).

With the recent recognition of defective hepatitis B viruses, new insight was gained into the structure-function relationship of different viral sequences. *In vitro* studies showed that the virus could retain the ability for helper-dependent packaging and reverse transcription despite the deletion of up to 80% of its genome (Horwich *et al.*, 1990). This suggested that large portions of the genome could be replaced with foreign genetic material. The hepatotropism and persistence (integration) were particularly attractive

properties for liver-directed gene transfer. Chang *et al.* (1991) introduced the chloramphenicol acetyltransferase (CAT) gene into duck hepatitis B virus genome in the place of the polymerase, surface, and pre-surface coding sequences. It was cotransfected with wild-type virus into an avian hepatoma cell line. Culture media containing high titers of the recombinant virus were used to infect primary duckling hepatocytes. Stable CAT gene expression was detected for at least 24 days after transfection (Chang *et al.*, 1991).

5. NON-VIRAL VECTORS

In order to effect expression of the oligonucleotide or polynucleotide sequences of the present invention, the expression construct must be delivered into a cell.

This delivery may be accomplished *in vitro*, as in laboratory procedures for transforming cells lines, or *in vivo* or *ex vivo*, as in the treatment of certain disease states. As described above, one preferred mechanism for delivery is *via* viral infection where the expression construct is encapsulated in an infectious viral particle.

Once the expression construct has been delivered into the cell the nucleic acid encoding the desired oligonucleotide or polynucleotide sequences may be positioned and expressed at different sites. In certain embodiments, the nucleic acid encoding the construct may be stably integrated into the genome of the cell. This integration may be in the specific location and orientation *via* homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the nucleic acid may be stably maintained in the cell as a separate, episomal segment of DNA. Such nucleic acid segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. How the expression construct is delivered to a cell and where in the cell the nucleic acid remains is dependent on the type of expression construct employed.

In certain embodiments of the invention, the expression construct comprising one or more oligonucleotide or polynucleotide sequences may simply consist of naked recombinant DNA or plasmids. Transfer of the construct may be performed by any of the methods mentioned above which physically or chemically permeabilize the cell

membrane. This is particularly applicable for transfer *in vitro* but it may be applied to *in vivo* use as well. Dubensky *et al.* (1984) successfully injected polyomavirus DNA in the form of calcium phosphate precipitates into liver and spleen of adult and newborn mice demonstrating active viral replication and acute infection. Benvenisty and Reshef (1986) also demonstrated that direct intraperitoneal injection of calcium phosphate-precipitated plasmids results in expression of the transfected genes. It is envisioned that DNA encoding a gene of interest may also be transferred in a similar manner *in vivo* and express the gene product.

Another embodiment of the invention for transferring a naked DNA expression construct into cells may involve particle bombardment. This method depends on the ability to accelerate DNA-coated microprojectiles to a high velocity allowing them to pierce cell membranes and enter cells without killing them (Klein *et al.*, 1987). Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang *et al.*, 1990). The microprojectiles used have consisted of biologically inert substances such as tungsten or gold beads.

Selected organs including the liver, skin, and muscle tissue of rats and mice have been bombarded *in vivo* (Yang *et al.*, 1990; Zelenin *et al.*, 1991). This may require surgical exposure of the tissue or cells, to eliminate any intervening tissue between the gun and the target organ, *i.e. ex vivo* treatment. Again, DNA encoding a particular gene may be delivered *via* this method and still be incorporated by the present invention.

ANTISENSE OLIGONUCLEOTIDES

The end result of the flow of genetic information is the synthesis of protein. DNA is transcribed by polymerases into messenger RNA and translated on the ribosome to yield a folded, functional protein. Thus there are several steps along the route where protein synthesis can be inhibited. The native DNA segment coding for a polypeptide described herein, as all such mammalian DNA strands, has two strands: a sense strand and an antisense strand held together by hydrogen bonding. The messenger RNA coding for

polypeptide has the same nucleotide sequence as the sense DNA strand except that the DNA thymidine is replaced by uridine. Thus, synthetic antisense nucleotide sequences will bind to a mRNA and inhibit expression of the protein encoded by that mRNA.

The targeting of antisense oligonucleotides to mRNA is thus one mechanism
5 to shut down protein synthesis, and, consequently, represents a powerful and targeted therapeutic approach. For example, the synthesis of polygalacturonase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829, each specifically incorporated herein by reference in its entirety). Further, examples of antisense
10 inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal GABA_A receptor and human EGF (Jaskulski *et al.*, 1988; Vasanthakumar and Ahmed, 1989; Peris *et al.*, 1998; U. S. Patent 5,801,154; U. S. Patent 5,789,573; U. S. Patent 5,718,709 and U. S. Patent 5,610,288, each specifically incorporated herein by reference in its entirety). Antisense
15 constructs have also been described that inhibit and can be used to treat a variety of abnormal cellular proliferations, *e.g.* cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683, each specifically incorporated herein by reference in its entirety).

Therefore, in exemplary embodiments, the invention provides
20 oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a
25 phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein.

Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence (*i.e.* in these illustrative examples the rat and human sequences) and determination of secondary structure, T_m , binding energy, relative stability, and antisense compositions were selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell.

Highly preferred target regions of the mRNA, are those which are at or near the AUG translation initiation codon, and those sequences which were substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations were performed using v.4 of the OLIGO primer analysis software (Rychlik, 1997) and the BLASTN 2.0.5 algorithm software (Altschul *et al.*, 1997).

The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, 1997). It has been demonstrated that several molecules of the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%). Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane (Morris *et al.*, 1997).

RIBOZYMES

Although proteins traditionally have been used for catalysis of nucleic acids, another class of macromolecules has emerged as useful in this endeavor. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, 1987; Gerlach *et al.*, 1987; Forster and Symons, 1987). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et*

al., 1981; Michel and Westhof, 1990; Reinhold-Hurek and Shub, 1992). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

5 Ribozyme catalysis has primarily been observed as part of sequence-specific cleavage/ligation reactions involving nucleic acids (Joyce, 1989; Cech *et al.*, 1981). For example, U. S. Patent No. 5,354,855 (specifically incorporated herein by reference) reports that certain ribozymes can act as endonucleases with a sequence specificity greater than that of known ribonucleases and approaching that of the DNA restriction enzymes. Thus, 10 sequence-specific ribozyme-mediated inhibition of gene expression may be particularly suited to therapeutic applications (Scanlon *et al.*, 1991; Sarver *et al.*, 1990). Recently, it was reported that ribozymes elicited genetic changes in some cells lines to which they were applied; the altered genes included the oncogenes *H-ras*, *c-fos* and genes of HIV. Most of this work involved the modification of a target mRNA, based on a specific mutant codon 15 that is cleaved by a specific ribozyme.

 Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through 20 the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an 25 encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

 The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds to

a nucleic acid target to block its translation) since the concentration of ribozyme necessary to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the
5 ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf *et al.*, 1992). Thus, the specificity of action
10 of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are
15 described by Rossi *et al.* (1992). Examples of hairpin motifs are described by Hampel *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz (1989), Hampel *et al.* (1990) and U. S. Patent 5,631,359 (specifically incorporated herein by reference). An example of the hepatitis δ virus motif is described by Perrotta and Been (1992); an example of the RNaseP motif is described by Guerrier-Takada *et al.* (1983); Neurospora VS RNA
20 ribozyme motif is described by Collins (Saville and Collins, 1990; Saville and Collins, 1991; Collins and Olive, 1993); and an example of the Group I intron is described in (U. S. Patent 4,987,071, specifically incorporated herein by reference). All that is important in an enzymatic nucleic acid molecule of this invention is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it
25 have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

In certain embodiments, it may be important to produce enzymatic cleaving agents which exhibit a high degree of specificity for the RNA of a desired target, such as

one of the sequences disclosed herein. The enzymatic nucleic acid molecule is preferably targeted to a highly conserved sequence region of a target mRNA. Such enzymatic nucleic acid molecules can be delivered exogenously to specific cells as required. Alternatively, the ribozymes can be expressed from DNA or RNA vectors that are delivered to specific
5 cells.

Small enzymatic nucleic acid motifs (*e.g.*, of the hammerhead or the hairpin structure) may also be used for exogenous delivery. The simple structure of these molecules increases the ability of the enzymatic nucleic acid to invade targeted regions of the mRNA structure. Alternatively, catalytic RNA molecules can be expressed within cells
10 from eukaryotic promoters (*e.g.*, Scanlon *et al.*, 1991; Kashani-Sabet *et al.*, 1992; Dropulic *et al.*, 1992; Weerasinghe *et al.*, 1991; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Sarver *et al.*, 1990). Those skilled in the art realize that any ribozyme can be expressed in eukaryotic cells from the appropriate DNA vector. The activity of such ribozymes can be augmented by their release from the primary transcript by a second ribozyme (Int. Pat. Appl. Publ. No.
15 WO 93/23569, and Int. Pat. Appl. Publ. No. WO 94/02595, both hereby incorporated by reference; Ohkawa *et al.*, 1992; Taira *et al.*, 1991; and Ventura *et al.*, 1993).

Ribozymes may be added directly, or can be complexed with cationic lipids, lipid complexes, packaged within liposomes, or otherwise delivered to target cells. The RNA or RNA complexes can be locally administered to relevant tissues *ex vivo*, or *in vivo*
20 through injection, aerosol inhalation, infusion pump or stent, with or without their incorporation in biopolymers.

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such
25 ribozymes can also be optimized for delivery. While specific examples are provided, those in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

Hammerhead or hairpin ribozymes may be individually analyzed by computer folding (Jaeger *et al.*, 1989) to assess whether the ribozyme sequences fold into

the appropriate secondary structure. Those ribozymes with unfavorable intramolecular interactions between the binding arms and the catalytic core are eliminated from consideration. Varying binding arm lengths can be chosen to optimize activity. Generally, at least 5 or so bases on each arm are able to bind to, or otherwise interact with, the target

5 RNA.

Ribozymes of the hammerhead or hairpin motif may be designed to anneal to various sites in the mRNA message, and can be chemically synthesized. The method of synthesis used follows the procedure for normal RNA synthesis as described in Usman *et al.* (1987) and in Scaringe *et al.* (1990) and makes use of common nucleic acid

10 protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. Average stepwise coupling yields are typically >98%. Hairpin ribozymes may be synthesized in two parts and annealed to reconstruct an active ribozyme (Chowrira and Burke, 1992). Ribozymes may be modified extensively to enhance stability by modification with nuclease resistant groups, for example, 2'-amino, 2'-

15 C-allyl, 2'-fluoro, 2'-o-methyl, 2'-H (for a review see *e.g.*, Usman and Cedergren, 1992). Ribozymes may be purified by gel electrophoresis using general methods or by high pressure liquid chromatography and resuspended in water.

Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their

20 degradation by serum ribonucleases (see *e.g.*, Int. Pat. Appl. Publ. No. WO 92/07065; Perrault *et al.*, 1990; Pieken *et al.*, 1991; Usman and Cedergren, 1992; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of

25 enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be

administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stent. Other routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector. Transcription of the ribozyme sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, *etc.*) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells (Elroy-Stein and Moss, 1990; Gao and Huang, 1993; Lieber *et al.*, 1993; Zhou *et al.*, 1990). Ribozymes expressed from such promoters can function in mammalian cells (*e.g.* Kashani-Saber *et al.*, 1992; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Yu *et al.*, 1993; L'Huillier *et al.*, 1992; Lisiewicz *et al.*, 1993). Such transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

Ribozymes may be used as diagnostic tools to examine genetic drift and mutations within diseased cells. They can also be used to assess levels of the target RNA molecule. The close relationship between ribozyme activity and the structure of the target RNA allows the detection of mutations in any region of the molecule which alters the base-pairing and three-dimensional structure of the target RNA. By using multiple ribozymes, one may map nucleotide changes which are important to RNA structure and function *in vitro*, as well as in cells and tissues. Cleavage of target RNAs with ribozymes may be used to inhibit gene expression and define the role (essentially) of specified gene products in the progression of disease. In this manner, other genetic targets may be defined as important mediators of the disease. These studies will lead to better treatment of the disease progression by affording the possibility of combinational therapies (e.g., multiple ribozymes targeted to different genes, ribozymes coupled with known small molecule inhibitors, or intermittent treatment with combinations of ribozymes and/or other chemical or biological molecules). Other *in vitro* uses of ribozymes are well known in the art, and include detection of the presence of mRNA associated with an IL-5 related condition. Such RNA is detected by determining the presence of a cleavage product after treatment with a ribozyme using standard methodology.

PEPTIDE NUCLEIC ACIDS

In certain embodiments, the inventors contemplate the use of peptide nucleic acids (PNAs) in the practice of the methods of the invention. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and Nielsen, 1997). PNA is able to be utilized in a number methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (1997) and is incorporated herein by reference. As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter,

PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, 1991; Hanvey *et al.*, 1992; Hyrup and Nielsen, 1996; Neilsen, 1996). This chemistry has three important consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc (Dueholm *et al.*, 1994) or Fmoc (Thomson *et al.*, 1995) protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used (Christensen *et al.*, 1995).

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this difficulty, it is suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed by the purification of PNAs by reverse-phase high-pressure liquid chromatography (Norton *et al.*, 1995) providing yields and purity of product similar to those observed during the synthesis of peptides.

67

functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (Norton *et al.*, 1995; Haaime *et al.*, 1996; Stetsenko *et al.*, 1996; Petersen *et al.*, 1995; Ulmann *et al.*, 1996; Koch *et al.*, 1995; Orum *et al.*, 1995; Footer *et al.*, 1996; Griffith *et al.*, 1995; Kremisky *et al.*, 1996; Pardridge *et al.*, 1995; Boffa *et al.*, 1995; Landsdorp *et al.*, 1996; Gambacorti-Passerini *et al.*, 1996; Armitage *et al.*, 1997; Seeger *et al.*, 1997; Ruskowski *et al.*, 1997). U.S. Patent No. 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

In contrast to DNA and RNA, which contain negatively charged linkages, the PNA backbone is neutral. In spite of this dramatic alteration, PNAs recognize complementary DNA and RNA by Watson-Crick pairing (Egholm *et al.*, 1993), validating the initial modeling by Nielsen *et al.* (1991). PNAs lack 3' to 5' polarity and can bind in either parallel or antiparallel fashion, with the antiparallel mode being preferred (Egholm *et al.*, 1993).

Hybridization of DNA oligonucleotides to DNA and RNA is destabilized by electrostatic repulsion between the negatively charged phosphate backbones of the complementary strands. By contrast, the absence of charge repulsion in PNA-DNA or PNA-RNA duplexes increases the melting temperature (T_m) and reduces the dependence of T_m on the concentration of mono- or divalent cations (Nielsen *et al.*, 1991). The enhanced rate and affinity of hybridization are significant because they are responsible for the surprising ability of PNAs to perform strand invasion of complementary sequences within relaxed double-stranded DNA. In addition, the efficient hybridization at inverted repeats suggests that PNAs can recognize secondary structure effectively within double-stranded DNA. Enhanced recognition also occurs with PNAs immobilized on surfaces, and Wang *et al.* have shown that support-bound PNAs can be used to detect hybridization events (Wang *et al.*, 1996).

One might expect that tight binding of PNAs to complementary sequences would also increase binding to similar (but not identical) sequences, reducing the sequence

5

10

15

20

Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (1993) and Jensen *et al.* (1997). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the relative binding kinetics and stoichiometry. Similar types of measurements were made by Jensen *et al.* using BIAcore™ technology.

Other applications of PNAs include use in DNA strand invasion (Nielsen *et al.*, 1991), antisense inhibition (Hanvey *et al.*, 1992), mutational analysis (Orum *et al.*, 1993), enhancers of transcription (Mollegaard *et al.*, 1994), nucleic acid purification (Orum *et al.*, 1995), isolation of transcriptionally active genes (Boffa *et al.*, 1995), blocking of transcription factor binding (Vickers *et al.*, 1995), genome cleavage (Veselkov *et al.*, 1996), biosensors (Wang *et al.*, 1996), *in situ* hybridization (Thisted *et al.*, 1996), and in a alternative to Southern blotting (Perry-O'Keefe, 1996).

POLYPEPTIDE COMPOSITIONS

The present invention, in other aspects, provides polypeptide compositions. Generally, a polypeptide of the invention will be an isolated polypeptide (or an epitope, variant, or active fragment thereof) derived from a mammalian species. Preferably, the polypeptide is encoded by a polynucleotide sequence disclosed herein or a sequence which hybridizes under moderately stringent conditions to a polynucleotide sequence disclosed herein. Alternatively, the polypeptide may be defined as a polypeptide which comprises a contiguous amino acid sequence from an amino acid sequence disclosed herein, or which polypeptide comprises an entire amino acid sequence disclosed herein.

In the present invention, a polypeptide composition is also understood to comprise one or more polypeptides that are immunologically reactive with antibodies generated against a polypeptide of the invention, particularly a polypeptide having the amino acid sequence disclosed in SEQ ID NO: 391 and 393, or to active fragments, or to variants or biological functional equivalents thereof.

Likewise, a polypeptide composition of the present invention is understood to comprise one or more polypeptides that are capable of eliciting antibodies that are

immunologically reactive with one or more polypeptides encoded by one or more contiguous nucleic acid sequences contained in SEQ ID NO: 217-390 and 392, or to active fragments, or to variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency.

5 Particularly illustrative polypeptides include the amino acid sequence disclosed in SEQ ID NO:391, 393, 395, 397, 421 and 425-427.

As used herein, an active fragment of a polypeptide includes a whole or a portion of a polypeptide which is modified by conventional techniques, *e.g.*, mutagenesis, or by addition, deletion, or substitution, but which active fragment exhibits substantially
10 the same structure function, antigenicity, etc., as a polypeptide as described herein.

In certain illustrative embodiments, the polypeptides of the invention will comprise at least an immunogenic portion of a lung tumor protein or a variant thereof, as described herein. As noted above, a "lung tumor protein" is a protein that is expressed by lung tumor cells. Proteins that are lung tumor proteins also react detectably within an
15 immunoassay (such as an ELISA) with antisera from a patient with lung cancer. Polypeptides as described herein may be of any length. Additional sequences derived from the native protein and/or heterologous sequences may be present, and such sequences may (but need not) possess further immunogenic or antigenic properties.

An "immunogenic portion," as used herein is a portion of a protein that is
20 recognized (*i.e.*, specifically bound) by a B-cell and/or T-cell surface antigen receptor. Such immunogenic portions generally comprise at least 5 amino acid residues, more preferably at least 10, and still more preferably at least 20 amino acid residues of a lung tumor protein or a variant thereof. Certain preferred immunogenic portions include peptides in which an N-terminal leader sequence and/or transmembrane domain have been
25 deleted. Other preferred immunogenic portions may contain a small N- and/or C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247

(Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well known techniques. An immunogenic portion of a native lung tumor protein is a portion that reacts with such antisera and/or T-cells at a level that is not substantially less than the reactivity of the full length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Such immunogenic portions may react within such assays at a level that is similar to or greater than the reactivity of the full length polypeptide. Such screens may generally be performed using methods well known to those of ordinary skill in the art, such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. For example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, ¹²⁵I-labeled Protein A.

As noted above, a composition may comprise a variant of a native lung tumor protein. A polypeptide "variant," as used herein, is a polypeptide that differs from a native lung tumor protein in one or more substitutions, deletions, additions and/or insertions, such that the immunogenicity of the polypeptide is not substantially diminished. In other words, the ability of a variant to react with antigen-specific antisera may be enhanced or unchanged, relative to the native protein, or may be diminished by less than 50%, and preferably less than 20%, relative to the native protein. Such variants may generally be identified by modifying one of the above polypeptide sequences and evaluating the reactivity of the modified polypeptide with antigen-specific antibodies or antisera as described herein. Preferred variants include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other preferred variants include variants in which a small portion (*e.g.*, 1-30

amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

Polypeptide variants encompassed by the present invention include those exhibiting at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%,
 5 97%, 98%, or 99% or more identity (determined as described above) to the polypeptides disclosed herein.

Preferably, a variant contains conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the
 10 secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Amino acid substitutions may generally be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids
 15 with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or
 20 alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the polypeptide.

25 As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (*e.g.*, poly-

His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

Polypeptides may be prepared using any of a variety of well known techniques. Recombinant polypeptides encoded by DNA sequences as described above
5 may be readily prepared from the DNA sequences using any of a variety of expression vectors known to those of ordinary skill in the art. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast, and higher eukaryotic cells, such as mammalian cells and plant
10 cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. Supernatants from suitable host/vector systems which secrete recombinant protein or polypeptide into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be applied to a suitable purification matrix such as an affinity matrix or an ion exchange resin. Finally,
15 one or more reverse phase HPLC steps can be employed to further purify a recombinant polypeptide.

Portions and other variants having less than about 100 amino acids, and generally less than about 50 amino acids, may also be generated by synthetic means, using techniques well known to those of ordinary skill in the art. For example, such polypeptides
20 may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. See Merrifield, *J. Am. Chem. Soc.* 85:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may
25 be operated according to the manufacturer's instructions.

Within certain specific embodiments, a polypeptide may be a fusion protein that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological

5

10

20

generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

Fusion proteins are also provided. Such proteins comprise a polypeptide as described herein together with an unrelated immunogenic protein. Preferably the immunogenic protein is capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (*see, for example, Stoute et al. New Engl. J. Med.*, 336:86-91, 1997).

Within preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium *Haemophilus influenza B* (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (*e.g.*, the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemagglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the *LytA* gene; *Gene* 43:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-

terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (*see Biotechnology 10:795-798, 1992*). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion protein. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

In general, polypeptides (including fusion proteins) and polynucleotides as described herein are isolated. An "isolated" polypeptide or polynucleotide is one that is removed from its original environment. For example, a naturally-occurring protein is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure. A polynucleotide is considered to be isolated if, for example, it is cloned into a vector that is not a part of the natural environment.

BINDING AGENTS

The present invention further provides agents, such as antibodies and antigen-binding fragments thereof, that specifically bind to a lung tumor protein. As used herein, an antibody, or antigen-binding fragment thereof, is said to "specifically bind" to a lung tumor protein if it reacts at a detectable level (within, for example, an ELISA) with a lung tumor protein, and does not react detectably with unrelated proteins under similar conditions. As used herein, "binding" refers to a noncovalent association between two separate molecules such that a complex is formed. The ability to bind may be evaluated by, for example, determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by the product of the component concentrations. In general, two compounds are said to "bind," in the context of the present invention, when the binding constant for complex

formation exceeds about 10^3 L/mol. The binding constant may be determined using methods well known in the art.

Binding agents may be further capable of differentiating between patients with and without a cancer, such as lung cancer, using the representative assays provided herein. In other words, antibodies or other binding agents that bind to a lung tumor protein will generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, and will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. It will be apparent that a statistically significant number of samples with and without the disease should be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as

bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

Within certain embodiments, the use of antigen-binding fragments of antibodies may be preferred. Such fragments include Fab fragments, which may be prepared using standard techniques. Briefly, immunoglobulins may be purified from rabbit serum by affinity chromatography on Protein A bead columns (Harlow and Lane, 5 *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988) and digested by papain to yield Fab and Fc fragments. The Fab and Fc fragments may be separated by affinity chromatography on protein A bead columns.

Monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, 10 differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y , ^{123}I , ^{125}I , ^{131}I , ^{186}Re , ^{188}Re , ^{211}At , and ^{212}Bi . Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria toxin, cholera toxin, gelonin, *Pseudomonas* exotoxin, *Shigella* toxin, and pokeweed 15 antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or 20 sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an 25 antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl
5 groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell *et al.*

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of
10 different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter *et al.*), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn *et al.*), by serum complement-
15 mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell *et al.*), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler *et al.*).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody.
20 Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent
25 bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato *et al.*), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih *et al.*). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide

agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison *et al.* discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous, intramuscular, subcutaneous or in the bed of a resected tumor. It will be evident that the precise dose of the antibody/immunoconjugate will vary depending upon the antibody used, the antigen density on the tumor, and the rate of clearance of the antibody.

T CELLS

Immunotherapeutic compositions may also, or alternatively, comprise T cells specific for a lung tumor protein. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a lung tumor polypeptide, polynucleotide encoding a lung tumor polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide. Preferably, a lung tumor polypeptide or polynucleotide is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a lung tumor polypeptide if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the

polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such

5 assays may be performed, for example, as described in Chen *et al.*, *Cancer Res.* 54:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (*e.g.*, by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into

10 DNA). Contact with a lung tumor polypeptide (100 ng/ml - 100 µg/ml, preferably 200 ng/ml - 25 µg/ml) for 3 - 7 days should result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.*, TNF or IFN-γ) is indicative of T cell activation (*see* Coligan *et al.*,

15 Current Protocols in Immunology, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a lung tumor polypeptide, polynucleotide or polypeptide-expressing APC may be CD4⁺ and/or CD8⁺. Lung tumor protein-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered

20 to the patient following stimulation and expansion.

For therapeutic purposes, CD4⁺ or CD8⁺ T cells that proliferate in response to a lung tumor polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a lung tumor polypeptide, or a short

25 peptide corresponding to an immunogenic portion of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a lung tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of a lung tumor protein can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

PHARMACEUTICAL COMPOSITIONS

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions disclosed herein in pharmaceutically-acceptable solutions for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

It will also be understood that, if desired, the nucleic acid segment, RNA, DNA or PNA compositions that express a polypeptide as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or DNA compositions.

Formulation of pharmaceutically-acceptable excipients and carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation.

1. ORAL DELIVERY

In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (Mathiowitz *et al.*, 1997; Hwang *et al.*, 1998; U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451, each specifically
5 incorporated herein by reference in its entirety). The tablets, troches, pills, capsules and the like may also contain the following: a binder, as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent,
10 such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar, or both. A syrup of elixir may contain the active compound sucrose as a sweetening agent
15 methyl and propylparabens as preservatives, a dye and flavoring, such as cherry or orange flavor. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and formulations.

20 Typically, these formulations may contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared in such a way
25 that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. For example, a mouthwash may be prepared incorporating the active ingredient in the required amount in an appropriate solvent, such as a sodium borate solution (Dobell's Solution). Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water, binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

2. INJECTABLE DELIVERY

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even intraperitoneally as described in U. S. Patent 5,543,158; U. S. Patent 5,641,515 and U. S. Patent 5,399,363 (each specifically incorporated herein by reference in its entirety). Solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (U. S. Patent 5,466,468, specifically incorporated herein by reference in its entirety). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing,

for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of
5 surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying
10 absorption, for example, aluminum monostearate and gelatin.

For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In
15 this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will
20 necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject. Moreover, for human administration, preparations should meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

25 Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other

ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

5 The compositions disclosed herein may be formulated in a neutral or salt form. Pharmaceutically-acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived
10 from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable
15 solutions, drug-release capsules, and the like.

As used herein, "carrier" includes any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art.
20 Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when
25 administered to a human. The preparation of an aqueous composition that contains a protein as an active ingredient is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified.

3. NASAL DELIVERY

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs *via* nasal aerosol sprays has been described *e.g.*, in U. S. Patent 5,756,353 and U. S. Patent 5,804,212 (each specifically incorporated herein by reference in its entirety). Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, 1998) and lysophosphatidyl-glycerol compounds (U. S. Patent 5,725,871, specifically incorporated herein by reference in its entirety) are also well-known in the pharmaceutical arts. Likewise, transmucosal drug delivery in the form of a polytetrafluoroethylene support matrix is described in U. S. Patent 5,780,045 (specifically incorporated herein by reference in its entirety).

4. LIPOSOME-, NANOCAPSULE-, AND MICROPARTICLE-MEDIATED DELIVERY

In certain embodiments, the inventors contemplate the use of liposomes, nanocapsules, microparticles, microspheres, lipid particles, vesicles, and the like, for the introduction of the compositions of the present invention into suitable host cells. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like.

Such formulations may be preferred for the introduction of pharmaceutically-acceptable formulations of the nucleic acids or constructs disclosed herein. The formation and use of liposomes is generally known to those of skill in the art (see for example, Couvreur *et al.*, 1977; Couvreur, 1988; Lasic, 1998; which describes the use of liposomes and nanocapsules in the targeted antibiotic therapy for intracellular bacterial infections and diseases). Recently, liposomes were developed with improved serum stability and circulation half-times (Gabizon and Papahadjopoulos, 1988; Allen and Choun, 1987; U. S. Patent 5,741,516, specifically incorporated herein by reference in its entirety). Further, various methods of liposome and liposome like preparations as potential drug carriers have been reviewed (Takakura, 1998; Chandran *et al.*, 1997; Margalit, 1995;

U. S. Patent 5,567,434; U. S. Patent 5,552,157; U. S. Patent 5,565,213; U. S. Patent 5,738,868 and U. S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally resistant to transfection by other procedures including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.*, 1990; Muller *et al.*, 1990). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, drugs (Heath and Martin, 1986; Heath *et al.*, 1986; Balazsovits *et al.*, 1989; Fresta and Puglisi, 1996), radiotherapeutic agents (Pikul *et al.*, 1987), enzymes (Imaizumi *et al.*, 1990a; Imaizumi *et al.*, 1990b), viruses (Faller and Baltimore, 1984), transcription factors and allosteric effectors (Nicolau and Gersonde, 1979) into a variety of cultured cell lines and animals. In addition, several successful clinical trails examining the effectiveness of liposome-mediated drug delivery have been completed (Lopez-Berestein *et al.*, 1985a; 1985b; Coune, 1988; Sculier *et al.*, 1988). Furthermore, several studies suggest that the use of liposomes is not associated with autoimmune responses, toxicity or gonadal localization after systemic delivery (Mori and Fukatsu, 1992).

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs). MLVs generally have diameters of from 25 nm to 4 μ m. Sonication of MLVs results in the formation of small unilamellar vesicles (SUVs) with diameters in the range of 200 to 500 Å, containing an aqueous solution in the core.

Liposomes bear resemblance to cellular membranes and are contemplated for use in connection with the present invention as carriers for the peptide compositions. They are widely suitable as both water- and lipid-soluble substances can be entrapped, *i.e.* in the aqueous spaces and within the bilayer itself, respectively. It is possible that the drug-bearing liposomes may even be employed for site-specific delivery of active agents by selectively modifying the liposomal formulation.

In addition to the teachings of Couvreur *et al.* (1977; 1988), the following information may be utilized in generating liposomal formulations. Phospholipids can form a variety of structures other than liposomes when dispersed in water, depending on the molar ratio of lipid to water. At low ratios the liposome is the preferred structure. The physical characteristics of liposomes depend on pH, ionic strength and the presence of divalent cations. Liposomes can show low permeability to ionic and polar substances, but at elevated temperatures undergo a phase transition which markedly alters their permeability. The phase transition involves a change from a closely packed, ordered structure, known as the gel state, to a loosely packed, less-ordered structure, known as the fluid state. This occurs at a characteristic phase-transition temperature and results in an increase in permeability to ions, sugars and drugs.

In addition to temperature, exposure to proteins can alter the permeability of liposomes. Certain soluble proteins, such as cytochrome c, bind, deform and penetrate the bilayer, thereby causing changes in permeability. Cholesterol inhibits this penetration of proteins, apparently by packing the phospholipids more tightly. It is contemplated that the most useful liposome formations for antibiotic and inhibitor delivery will contain cholesterol.

The ability to trap solutes varies between different types of liposomes. For example, MLVs are moderately efficient at trapping solutes, but SUVs are extremely inefficient. SUVs offer the advantage of homogeneity and reproducibility in size distribution, however, and a compromise between size and trapping efficiency is offered by large unilamellar vesicles (LUVs). These are prepared by ether evaporation and are three to four times more efficient at solute entrapment than MLVs.

In addition to liposome characteristics, an important determinant in entrapping compounds is the physicochemical properties of the compound itself. Polar compounds are trapped in the aqueous spaces and nonpolar compounds bind to the lipid bilayer of the vesicle. Polar compounds are released through permeation or when the bilayer is broken, but nonpolar compounds remain affiliated with the bilayer unless it is

disrupted by temperature or exposure to lipoproteins. Both types show maximum efflux rates at the phase transition temperature.

Liposomes interact with cells *via* four different mechanisms: endocytosis by phagocytic cells of the reticuloendothelial system such as macrophages and neutrophils; 5 adsorption to the cell surface, either by nonspecific weak hydrophobic or electrostatic forces, or by specific interactions with cell-surface components; fusion with the plasma cell membrane by insertion of the lipid bilayer of the liposome into the plasma membrane, with simultaneous release of liposomal contents into the cytoplasm; and by transfer of liposomal lipids to cellular or subcellular membranes, or vice versa, without any association of the 10 liposome contents. It often is difficult to determine which mechanism is operative and more than one may operate at the same time.

The fate and disposition of intravenously injected liposomes depend on their physical properties, such as size, fluidity, and surface charge. They may persist in tissues for h or days, depending on their composition, and half lives in the blood range from min to 15 several h. Larger liposomes, such as MLVs and LUVs, are taken up rapidly by phagocytic cells of the reticuloendothelial system, but physiology of the circulatory system restrains the exit of such large species at most sites. They can exit only in places where large openings or pores exist in the capillary endothelium, such as the sinusoids of the liver or spleen. Thus, these organs are the predominate site of uptake. On the other hand, SUVs 20 show a broader tissue distribution but still are sequestered highly in the liver and spleen. In general, this *in vivo* behavior limits the potential targeting of liposomes to only those organs and tissues accessible to their large size. These include the blood, liver, spleen, bone marrow, and lymphoid organs.

Targeting is generally not a limitation in terms of the present invention. 25 However, should specific targeting be desired, methods are available for this to be accomplished. Antibodies may be used to bind to the liposome surface and to direct the antibody and its drug contents to specific antigenic receptors located on a particular cell-type surface. Carbohydrate determinants (glycoprotein or glycolipid cell-surface components that play a role in cell-cell recognition, interaction and adhesion) may also be

used as recognition sites as they have potential in directing liposomes to particular cell types. Mostly, it is contemplated that intravenous injection of liposomal preparations would be used, but other routes of administration are also conceivable.

Alternatively, the invention provides for pharmaceutically-acceptable
5 nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (Henry-Michelland *et al.*, 1987; Quintanar-Guerrero *et al.*, 1998; Douglas *et al.*, 1987). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-
10 cyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present invention. Such particles may be easily made, as described (Couvreux *et al.*, 1980; 1988; zur Muhlen *et al.*, 1998; Zambaux *et al.* 1998; Pinto-Alphandry *et al.*, 1995 and U. S. Patent 5,145,684, specifically incorporated herein by reference in its entirety).

IMMUNOGENIC COMPOSITIONS

15 In certain preferred embodiments of the present invention, immunogenic compositions, or vaccines, are provided. The immunogenic compositions will generally comprise one or more pharmaceutical compositions, such as those discussed above, in combination with an immunostimulant. An immunostimulant may be any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an
20 exogenous antigen. Examples of immunostimulants include adjuvants, biodegradable microspheres (*e.g.*, polylactic galactide) and liposomes (into which the compound is incorporated; *see e.g.*, Fullerton, U.S. Patent No. 4,235,877). Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Pharmaceutical
25 compositions and immunogenic compositions within the scope of the present invention may also contain other compounds, which may be biologically active or inactive. For example, one or more immunogenic portions of other tumor antigens may be present, either incorporated into a fusion polypeptide or as a separate compound, within the composition.

Illustrative immunogenic compositions may contain DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, bacteria and viral expression systems. Numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, 1998, and references cited therein. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope. In a preferred embodiment, the DNA may be introduced using a viral expression system (*e.g.*, vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are disclosed, for example, in Fisher-Hoch *et al.*, *Proc. Natl. Acad. Sci. USA* 86:317-321, 1989; Flexner *et al.*, *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner *et al.*, *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques* 6:616-627, 1988; Rosenfeld *et al.*, *Science* 252:431-434, 1991; Kolls *et al.*, *Proc. Natl. Acad. Sci. USA* 91:215-219, 1994; Kass-Eisler *et al.*, *Proc. Natl. Acad. Sci. USA* 90:11498-11502, 1993; Guzman *et al.*, *Circulation* 88:2838-2848, 1993; and Guzman *et al.*, *Cir. Res.* 73:1202-1207, 1993. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer *et al.*, *Science* 259:1745-1749, 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells. It will be apparent that an immunogenic composition may comprise both a polynucleotide and a polypeptide component. Such immunogenic compositions may provide for an enhanced immune response.

It will be apparent that an immunogenic composition may contain pharmaceutically acceptable salts of the polynucleotides and polypeptides provided herein. Such salts may be prepared from pharmaceutically acceptable non-toxic bases, including organic bases (*e.g.*, salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (*e.g.*, sodium, potassium, lithium, ammonium, calcium and magnesium salts).

While any suitable carrier known to those of ordinary skill in the art may be employed in the compositions of this invention, the type of carrier will vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (*e.g.*, polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344 and 5,942,252. One may also employ a carrier comprising the particulate-protein complexes described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

Such compositions may also comprise buffers (*e.g.*, neutral buffered saline or phosphate buffered saline), carbohydrates (*e.g.*, glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (*e.g.*, aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively,

compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known technology.

Any of a variety of immunostimulants may be employed in the immunogenic compositions of this invention. For example, an adjuvant may be included.

- 5 Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Suitable adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and
- 10 Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -
- 15 12, may also be used as adjuvants.

- Within the immunogenic compositions provided herein, the adjuvant composition is preferably designed to induce an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.*, IFN- γ , TNF α , IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In
- 20 contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of an immunogenic composition as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent
- 25 than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

Preferred adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-

acylated monophosphoryl lipid A (3D-MPL), together with an aluminum salt. MPL adjuvants are available from Corixa Corporation (Seattle, WA; *see* US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response.

- 5 Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato *et al.*, *Science* 273:352, 1996. Another preferred adjuvant is a saponin, preferably QS21 (Aquila Biopharmaceuticals Inc., Framingham, MA), which may be used alone or in combination with other adjuvants. For
- 10 example, an enhanced system involves the combination of a monophosphoryl lipid A and saponin derivative, such as the combination of QS21 and 3D-MPL as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. A particularly potent adjuvant formulation involving QS21, 3D-
- 15 MPL and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

- Other preferred adjuvants include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (*e.g.*, SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and
- 20 other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties.

- Any immunogenic composition provided herein may be prepared using well known methods that result in a combination of antigen, immune response enhancer and a
- 25 suitable carrier or excipient. The compositions described herein may be administered as part of a sustained release formulation (*i.e.*, a formulation such as a capsule, sponge or gel (composed of polysaccharides, for example) that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology (*see, e.g.*, Coombes *et al.*, *Vaccine* 14:1429-1438, 1996) and administered by,

for example, oral, rectal or subcutaneous implantation, or by implantation at the desired target site. Sustained-release formulations may contain a polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane.

5 Carriers for use within such formulations are biocompatible, and may also be biodegradable; preferably the formulation provides a relatively constant level of active component release. Such carriers include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (*e.g.*, a
10 cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (*see e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the
15 condition to be treated or prevented.

Any of a variety of delivery vehicles may be employed within pharmaceutical compositions and immunogenic compositions to facilitate production of an antigen-specific immune response that targets tumor cells. Delivery vehicles include antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes
20 and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs,
25 including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature* 392:245-251, 1998) and have been shown to be

effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (see Timmerman and Levy, *Ann. Rev. Med.* 50:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and
 5 present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be
 10 used within an immunogenic composition (see Zitvogel *et al.*, *Nature Med.* 4:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells
 15 may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNF α to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNF α , CD40 ligand, LPS, flt3 ligand and/or other
 20 compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC
 25 with a high capacity for antigen uptake and processing, which correlates with the high expression of Fc γ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion

molecules (*e.g.*, CD54 and CD11) and costimulatory molecules (*e.g.*, CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide encoding a lung tumor protein (or portion or other variant thereof) such that the lung tumor polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a composition comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi *et al.*, *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the lung tumor polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (*e.g.*, vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (*e.g.*, a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

Immunogenic compositions and pharmaceutical compositions may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are preferably hermetically sealed to preserve sterility of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a immunogenic composition or pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

CANCER THERAPY

In further aspects of the present invention, the compositions described herein may be used for immunotherapy of cancer, such as lung cancer. Within such methods, pharmaceutical compositions and immunogenic compositions are typically administered to a patient. As used herein, a “patient” refers to any warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions and immunogenic compositions may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. A cancer may be diagnosed using criteria generally accepted in the art, including the presence of a malignant tumor. Pharmaceutical compositions and immunogenic compositions may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. Administration may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as CD8⁺ cytotoxic T lymphocytes and CD4⁺ T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other

vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow *in vivo* and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see*, for example, Cheever *et al.*, *Immunological Reviews* 157:177, 1997).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions

and immunogenic compositions may be administered by injection (*e.g.*, intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (*e.g.*, by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may

5 be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector

10 cells capable of killing the patient's tumor cells *in vitro*. Such immunogenic compositions should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free survival) in treated patients as compared to non-treated patients. In general, for pharmaceutical compositions and immunogenic compositions comprising one or more

15 polypeptides, the amount of each polypeptide present in a dose ranges from about 25 μ g to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit.

20 Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a lung tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine

25 assays, which may be performed using samples obtained from a patient before and after treatment.

CANCER DETECTION AND DIAGNOSIS

In general, a cancer may be detected in a patient based on the presence of one or more lung tumor proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as lung cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, a lung tumor sequence should be present at a level that is at least three fold higher in tumor tissue than in normal tissue

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the

labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length lung tumor proteins and portions thereof to which the binding agent binds, as described above.

5 The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic
10 particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which
15 may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1
20 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μ g, and preferably about 100 ng to about 1 μ g, is sufficient to immobilize an adequate amount of binding agent.

 Covalent attachment of binding agent to a solid support may generally be
25 achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group

on the support with an amine and an active hydrogen on the binding partner (*see, e.g.*, Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with lung cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound
 5 detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group
 10 (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as lung cancer, the signal detected from the reporter group that remains bound to the solid support is generally
 15 compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate
 20 preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett *et al.*, *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value
 25 for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the

false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1 μ g, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use lung tumor polypeptides to detect antibodies that bind to such polypeptides in a biological sample.

The detection of such lung tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a lung tumor protein in a biological sample. Within certain methods, a biological sample comprising CD4⁺ and/or CD8⁺ T cells isolated from a patient is incubated with a lung tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (*e.g.*, 5 - 25 µg/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of lung tumor polypeptide to serve as a control. For CD4⁺ T cells, activation is preferably detected by evaluating proliferation of the T cells. For CD8⁺ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a lung tumor protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a lung tumor cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the lung tumor protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a lung tumor protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a lung tumor protein that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence recited in SEQ ID NOs:217-390, 392, 394, 396, 398-420 and 422-424. Techniques for both PCR based assays and hybridization assays are well known in the art (*see*, for example, Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.*, 51:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

In another embodiment, the compositions described herein may be used as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed

as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

5 Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

10 As noted above, to improve sensitivity, multiple lung tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein markers may be based on routine experiments to determine combinations that results in
15 optimal sensitivity. In addition, or alternatively, assays for tumor proteins provided herein may be combined with assays for other known tumor antigens.

DIAGNOSTIC KITS

 The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for
20 performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a lung tumor protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or
25 buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a lung tumor protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a lung tumor protein. Such an oligonucleotide may be used, for example, within

5 a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a lung tumor protein.

The following Examples are offered by way of illustration and not by way of limitation.

10

Example 1

PREPARATION OF LUNG TUMOR-SPECIFIC CDNA SEQUENCES USING DIFFERENTIAL DISPLAY RT-PCR

5 This example illustrates the preparation of cDNA molecules encoding lung tumor-specific polypeptides using a differential display screen.

 Tissue samples were prepared from lung tumor and normal tissue of a patient with lung cancer that was confirmed by pathology after removal of samples from the patient. Normal RNA and tumor RNA was extracted from the samples and mRNA was
10 isolated and converted into cDNA using a (dT)₁₂AG (SEQ ID NO: 47) anchored 3' primer. Differential display PCR was then executed using a randomly chosen primer (SEQ ID NO: 48). Amplification conditions were standard buffer containing 1.5 mM MgCl₂, 20 pmol of primer, 500 pmol dNTP and 1 unit of Taq DNA polymerase (Perkin-Elmer, Branchburg, NJ). Forty cycles of amplification were performed using 94 °C denaturation for 30
15 seconds, 42 °C annealing for 1 minute and 72 °C extension for 30 seconds. Bands that were repeatedly observed to be specific to the RNA fingerprint pattern of the tumor were cut out of a silver stained gel, subcloned into the pGEM-T vector (Promega, Madison, WI) and sequenced. The isolated 3' sequences are provided in SEQ ID NO: 1-16.

 Comparison of these sequences to those in the public databases using the
20 BLASTN program, revealed no significant homologies to the sequences provided in SEQ ID NO: 1-11. To the best of the inventors' knowledge, none of the isolated DNA sequences have previously been shown to be expressed at a greater level in human lung tumor tissue than in normal lung tissue.

Example 2USE OF PATIENT SERA TO IDENTIFY DNA SEQUENCES ENCODING
LUNG TUMOR ANTIGENS

5

This example illustrates the isolation of cDNA sequences encoding lung tumor antigens by expression screening of lung tumor samples with autologous patient sera.

A human lung tumor directional cDNA expression library was constructed
10 employing the Lambda ZAP Express expression system (Stratagene, La Jolla, CA). Total RNA for the library was taken from a late SCID mouse passaged human squamous epithelial lung carcinoma and poly A⁺ RNA was isolated using the Message Maker kit (Gibco BRL, Gaithersburg, MD). The resulting library was screened using *E. coli*-absorbed autologous patient serum, as described in Sambrook et al., (*Molecular Cloning: A*
15 *Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989), with the secondary antibody being goat anti-human IgG-A-M (H + L) conjugated with alkaline phosphatase, developed with NBT/BCIP (Gibco BRL). Positive plaques expressing immunoreactive antigens were purified. Phagemid from the plaques was rescued and the nucleotide sequences of the clones was determined.

20 Fifteen clones were isolated, referred to hereinafter as LT86-1 – LT86-15. The isolated cDNA sequences for LT86-1 – LT86-8 and LT86-10 - LT86-15 are provided in SEQ ID NO: 17-24 and 26-31, respectively, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 32-39 and 41-46, respectively. The determined cDNA sequence for LT86-9 is provided in SEQ ID NO: 25, with the
25 corresponding predicted amino acid sequences from the 3' and 5' ends being provided in SEQ ID NO: 40 and 65, respectively. These sequences were compared to those in the gene bank as described above. Clones LT86-3, LT86-6 – LT86-9, LT86-11 – LT86-13 and LT86-15 (SEQ ID NO: 19, 22-25, 27-29 and 31, respectively) were found to show some homology to previously identified expressed sequence tags (ESTs), with clones LT86-6,

LT86-8, LT86-11, LT86-12 and LT86-15 appearing to be similar or identical to each other. Clone LT86-3 was found to show some homology with a human transcription repressor. Clones LT86-6, 8, 9, 11, 12 and 15 were found to show some homology to a yeast RNA Pol II transcription regulation mediator. Clone LT86-13 was found to show some
 5 homology with a *C. elegans* leucine aminopeptidase. Clone LT86-9 appears to contain two inserts, with the 5' sequence showing homology to the previously identified antisense sequence of interferon alpha-induced P27, and the 3' sequence being similar to LT86-6. Clone LT86-14 (SEQ ID NO: 30) was found to show some homology to the trithorax gene and has an "RGD" cell attachment sequence and a beta-Lactamase A site which functions
 10 in hydrolysis of penicillin. Clones LT86-1, LT86-2, LT86-4, LT86-5 and LT86-10 (SEQ ID NOS: 17, 18, 20, 21 and 26, respectively) were found to show homology to previously identified genes. A subsequently determined extended cDNA sequence for LT86-4 is provided in SEQ ID NO: 66, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 67.

15 Subsequent studies led to the isolation of five additional clones, referred to as LT86-20, LT86-21, LT86-22, LT86-26 and LT86-27. The determined 5' cDNA sequences for LT86-20, LT86-22, LT86-26 and LT86-27 are provided in SEQ ID NO: 68 and 70-72, respectively, with the determined 3' cDNA sequences for LT86-21 being provided in SEQ ID NO: 69. The corresponding predicted amino acid sequences for
 20 LT86-20, LT86-21, LT86-22, LT86-26 and LT86-27 are provided in SEQ ID NO: 73-77, respectively. LT86-22 and LT86-27 were found to be highly similar to each other. Comparison of these sequences to those in the gene bank as described above, revealed no significant homologies to LT86-22 and LT86-27. LT86-20, LT86-21 and LT86-26 were found to show homology to previously identified genes.

25 In further studies, a cDNA expression library was prepared using mRNA from a lung small cell carcinoma cell line in the lambda ZAP Express expression vector (Stratagene), and screened as described above, with a pool of two lung small cell carcinoma patient sera. The sera pool was adsorbed with *E. coli* lysate and human PBMC lysate was added to the serum to block antibody to proteins found in normal tissue.

Seventy-three clones were isolated. The determined cDNA sequences of these clones are provided in SEQ ID NO: 290-362. The sequences of SEQ ID NO: 289-292, 294, 296-297, 300, 302, 303, 305, 307-315, 317-320, 322-325, 327-332, 334, 335, 338-341, 343-352, 354-358, 360 and 362 were found to show some homology to previously isolated genes.

5 The sequences of SEQ ID NO: 293, 295, 298, 299, 301, 304, 306, 316, 321, 326, 333, 336, 337, 342, 353, 359 and 361 were found to show some homology to previously identified ESTs.

Example 3

10 USE OF MOUSE ANTISERA TO IDENTIFY DNA SEQUENCES ENCODING LUNG TUMOR ANTIGENS

This example illustrates the isolation of cDNA sequences encoding lung tumor antigens by screening of lung tumor cDNA libraries with mouse anti-tumor sera.

A directional cDNA lung tumor expression library was prepared as

15 described above in Example 2. Sera was obtained from SCID mice containing late passaged human squamous cell and adenocarcinoma tumors. These sera were pooled and injected into normal mice to produce anti-lung tumor serum. Approximately 200,000 PFUs were screened from the unamplified library using this antiserum. Using a goat anti-mouse IgG-A-M (H+L) alkaline phosphatase second antibody developed with NBT/BCIP (BRL

20 Labs.), approximately 40 positive plaques were identified. Phage was purified and phagemid excised for 9 clones with inserts in a pBK-CMV vector for expression in prokaryotic or eukaryotic cells.

The determined cDNA sequences for 7 of the isolated clones (hereinafter referred to as L86S-3, L86S-12, L86S-16, L86S-25, L86S-36, L86S-40 and L86S-46) are

25 provided in SEQ ID NO: 49-55, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 56-62, respectively. The 5' cDNA sequences for the remaining 2 clones (hereinafter referred to as L86S-30 and L86S-41) are provided in SEQ ID NO: 63 and 64. L86S-36 and L86S-46 were subsequently determined to represent the same gene. Comparison of these sequences with those in the public database as described

above, revealed no significant homologies to clones L86S-30, L86S-36 and L86S-46 (SEQ ID NO: 63, 53 and 55, respectively). L86S-16 (SEQ ID NO: 51) was found to show some homology to an EST previously identified in fetal lung and germ cell tumor. The remaining clones were found to show at least some degree of homology to previously identified human genes. Subsequently determined extended cDNA sequences for L86S-12, L86S-36 and L86S-46 are provided in SEQ ID NO: 78-80, respectively, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 81-83.

Subsequent studies led to the determination of 5' cDNA sequences for an additional nine clones, referred to as L86S-6, L86S-11, L86S-14, L86S-29, L86S-34, L86S-39, L86S-47, L86S-49 and L86S-51 (SEQ ID NO: 84-92, respectively). The corresponding predicted amino acid sequences are provided in SEQ ID NO: 93-101, respectively. L86S-30, L86S-39 and L86S-47 were found to be similar to each other. Comparison of these sequences with those in the gene bank as described above, revealed no significant homologies to L86S-14. L86S-29 was found to show some homology to a previously identified EST. L86S-6, L86S-11, L86S-34, L86S-39, L86S-47, L86S-49 and L86S-51 were found to show some homology to previously identified genes.

In further studies, a directional cDNA library was constructed using a Stratagene kit with a Lambda Zap Express vector. Total RNA for the library was isolated from two primary squamous lung tumors and poly A⁺ RNA was isolated using an oligo dT column. Antiserum was developed in normal mice using a pool of sera from three SCID mice implanted with human squamous lung carcinomas. Approximately 700,000 PFUs were screened from the unamplified library with *E. coli* absorbed mouse anti-SCID tumor serum. Positive plaques were identified as described above. Phage was purified and phagemid excised for 180 clones with inserts in a pBK-CMV vector for expression in prokaryotic or eukaryotic cells.

The determined cDNA sequences for 23 of the isolated clones are provided in SEQ ID NO: 126-148. Comparison of these sequences with those in the public database as described above revealed no significant homologies to the sequences of SEQ ID NO:

139 and 143-148. The sequences of SEQ ID NO: 126-138 and 140-142 were found to show homology to previously identified human polynucleotide sequences.

Example 4

5 USE OF MOUSE ANTISERA TO SCREEN LUNG TUMOR LIBRARIES PREPARED FROM SCID MICE

10 This example illustrates the isolation of cDNA sequences encoding lung tumor antigens by screening of lung tumor cDNA libraries prepared from SCID mice with mouse anti-tumor sera.

15 A directional cDNA lung tumor expression library was prepared using a Stratagene kit with a Lambda Zap Express vector. Total RNA for the library was taken from a late passaged lung adenocarcinoma grown in SCID mice. Poly A⁺ RNA was isolated using a Message Maker Kit (Gibco BRL). Sera was obtained from two SCID mice
20 implanted with lung adenocarcinomas. These sera were pooled and injected into normal mice to produce anti-lung tumor serum. Approximately 700,000 PFUs were screened from the unamplified library with *E. coli*-absorbed mouse anti-SCID tumor serum. Positive plaques were identified with a goat anti-mouse IgG-A-M (H+L) alkaline phosphatase second antibody developed with NBT/BCIP (Gibco BRL). Phage was purified and
25 phagemid excised for 100 clones with insert in a pBK-CMV vector for expression in prokaryotic or eukaryotic cells.

 The determined 5' cDNA sequences for 33 of the isolated clones are provided in SEQ ID NO: 149-181. The corresponding predicted amino acid sequences for SEQ ID NO: 149, 150, 152-154, 156-158 and 160-181 are provided in SEQ ID NO: 182,
25 183, 186, 188-193 and 194-215, respectively. The clone of SEQ ID NO: 151 (referred to as SAL-25) was found to contain two open reading frames (ORFs). The predicted amino acid sequences encoded by these ORFs are provided in SEQ ID NO: 184 and 185. The clone of SEQ ID NO: 153 (referred to as SAL-50) was found to contain two open reading frames encoding the predicted amino acid sequences of SEQ ID NO: 187 and 216.

Similarly, the clone of SEQ ID NO: 155 (referred to as SAL-66) was found to contain two open reading frames encoding the predicted amino acid sequences of SEQ ID NO: 189 and 190. Comparison of the isolated sequences with those in the public database revealed no significant homologies to the sequences of SEQ ID NO: 151, 153 and 154. The sequences of SEQ ID NO: 149, 152, 156, 157 and 158 were found to show some homology to previously isolated expressed sequence tags (ESTs). The sequences of SEQ ID NO: 150, 155 and 159-181 were found to show homology to sequences previously identified in humans.

Using the procedures described above, two directional cDNA libraries (referred to as LT46-90 and LT86-21) were prepared from two late passaged lung squamous carcinomas grown in SCID mice and screened with sera obtained from SCID mice implanted with human squamous lung carcinomas. The determined cDNA sequences for the isolated clones are provided in SEQ ID NO: 217-237 and 286-289. SEQ ID NO: 286 was found to be a longer sequence of LT4690-71 (SEQ ID NO: 237). Comparison of these sequences with those in the public databases revealed no known homologies to the sequences of SEQ ID NO: 219, 220, 225, 226, 287 and 288. The sequences of SEQ ID NO: 218, 221, 222 and 224 were found to show some homology to previously identified sequences of unknown function. The sequence of SEQ ID NO: 236 was found to show homology to a known mouse mRNA sequence. The sequences of SEQ ID NO: 217, 223, 227-237, 286 and 289 showed some homology to known human DNA and/or RNA sequences.

In further studies using the techniques described above, one of the cDNA libraries described above (LT86-21) was screened with *E. coli*-absorbed mouse anti-SCID tumor serum. This serum was obtained from normal mice immunized with a pool of 3 sera taken from SCID mice implanted with human squamous lung carcinomas. The determined
 5 cDNA sequences for the isolated clones are provided in SEQ ID NO: 238-285. Comparison of these sequences with those in the public databases revealed no significant homologies to the sequences of SEQ ID NO: 253, 260, 277 and 285. The sequences of SEQ ID NO: 249, 250, 256, 266, 276 and 282 were found to show some homology to previously isolated expressed sequence tags (ESTs). The sequences of SEQ ID NO: 238-248, 251, 252, 254,
 10 255, 257-259, 261-263, 265, 267-275, 278-281, 283 and 284 were found to show some homology to previously identified DNA or RNA sequences.

Full-length sequencing studies on antigen 2LT-128 (SEQ ID NO: 282) resulted in the isolation of the full-length cDNA sequence provided in SEQ ID NO: 392. This amino acid sequence encoded by this full-length cDNA sequence is provided in SEQ
 15 ID NO: 393. This antigen shows 20-fold over-expression in squamous cell carcinoma and 2.5-fold over-expression in lung adenocarcinoma. This gene has been described as a potential ras oncogene (Fenwick et al. *Science*, 287:869-873, 2000).

Extended sequence information was obtained for clones 2LT-3 (SEQ ID NO:238), 2LT-26 (SEQ ID NO:242), 2LT-57 (SEQ ID NO: 249), 2LT-58 (SEQ ID
 20 NO:250), 2LT-98 (SEQ ID NO:268) and 2LT-124 (SEQ ID NO:279). The extended cDNA sequences for these clones are set forth in SEQ ID NOs:428-433, respectively, encoding the polypeptide sequences set forth in SEQ ID NOs: 434-439, respectively.

Example 5

DETERMINATION OF TISSUE SPECIFICITY OF LUNG TUMOR POLYPEPTIDES

Using gene specific primers, mRNA expression levels for representative lung tumor polypeptides were examined in a variety of normal and tumor tissues using RT-PCR.

Briefly, total RNA was extracted from a variety of normal and tumor tissues using Trizol reagent. First strand synthesis was carried out using 2 µg of total RNA with SuperScript II reverse transcriptase (BRL Life Technologies) at 42 °C for one hour. The cDNA was then amplified by PCR with gene-specific primers. To ensure the semi-quantitative nature of the RT-PCR, β-actin was used as an internal control for each of the tissues examined. 1 µl of 1:30 dilution of cDNA was employed to enable the linear range amplification of the β-actin template and was sensitive enough to reflect the differences in the initial copy numbers. Using these conditions, the β-actin levels were determined for each reverse transcription reaction from each tissue. DNA contamination was minimized by DNase treatment and by assuring a negative PCR result when using first strand cDNA that was prepared without adding reverse transcriptase.

mRNA Expression levels were examined in five different types of tumor tissue (lung squamous tumor from 3 patients, lung adenocarcinoma, prostate tumor, colon tumor and lung tumor), and different normal tissues, including lung from four patients, prostate, brain, kidney, liver, ovary, skeletal muscle, skin, small intestine, myocardium, retina and testes. L86S-46 was found to be expressed at high levels in lung squamous tumor, colon tumor and prostate tumor, and was undetectable in the other tissues examined. L86S-5 was found to be expressed in the lung tumor samples and in 2 out of 4 normal lung samples, but not in the other normal or tumor tissues tested. L86S-16 was found to be expressed in all tissues except normal liver and normal stomach. Using real-time PCR, L86S-46 was found to be over-expressed in lung squamous tissue and normal tonsil, with expression being low or undetectable in all other tissues examined.

Example 6

ISOLATION OF DNA SEQUENCES ENCODING LUNG TUMOR ANTIGENS

DNA sequences encoding antigens potentially involved in squamous cell lung tumor formation were isolated as follows.

A lung tumor directional cDNA expression library was constructed employing the Lambda ZAP Express expression system (Stratagene, La Jolla, CA). Total RNA for the library was taken from a pool of two human squamous epithelial lung carcinomas and poly A⁺ RNA was isolated using oligo-dT cellulose (Gibco BRL, Gaithersburg, MD). Phagemid were rescued at random and the cDNA sequences of isolated clones were determined.

The determined cDNA sequence for the clone SLT-T1 is provided in SEQ ID NO: 102, with the determined 5' cDNA sequences for the clones SLT-T2, SLT-T3, SLT-T5, SLT-T7, SLT-T9, SLT-T10, SLT-T11 and SLT-T12 being provided in SEQ ID NO: 103-110, respectively. The corresponding predicted amino acid sequence for SLT-T1, SLT-T2, SLT-T3, SLT-T10 and SLT-T12 are provided in SEQ ID NO: 111-115, respectively. Comparison of the sequences for SLT-T2, SLT-T3, SLT-T5, SLT-T7, SLT-T9 and SLT-T11 with those in the public databases as described above, revealed no significant homologies. The sequences for SLT-T10 and SLT-T12 were found to show some homology to sequences previously identified in humans.

The sequence of SLT-T1 was determined to show some homology to a PAC clone of unknown protein function. The cDNA sequence of SLT-T1 (SEQ ID NO: 102) was found to contain a mutator (MUT) domain. Such domains are known to function in removal of damaged guanine from DNA that can cause A to G transversions (see, for example, el-Deiry, W.S., 1997 *Curr. Opin. Oncol.* 9:79-87; Okamoto, K. et al. 1996 *Int. J. Cancer* 65:437-41; Wu, C. et al. 1995 *Biochem. Biophys. Res. Commun.* 214:1239-45; Porter, D.W. et al. 1996 *Chem. Res. Toxicol.* 9:1375-81). SLT-T1 may thus be of use in the treatment, by gene therapy, of lung cancers caused by, or associated with, a disruption in DNA repair.

In further studies, DNA sequences encoding antigens potentially involved in adenocarcinoma lung tumor formation were isolated as follows. A human lung tumor directional cDNA expression library was constructed employing the Lambda ZAP Express expression system (Stratagene, La Jolla, CA). Total RNA for the library was taken from a late SCID mouse passaged human adenocarcinoma and poly A+ RNA was isolated using the Message Maker kit (Gibco BRL, Gaithersburg, MD). Phagemid were rescued at random and the cDNA sequences of isolated clones were determined.

The determined 5' cDNA sequences for five isolated clones (referred to as SALT-T3, SALT-T4, SALT-T7, SALT-T8, and SALT-T9) are provided in SEQ ID NO: 116-120, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 121-125. SALT-T3 was found to show 98% identity to the previously identified human transducin-like enhancer protein TLE2. SALT-T4 appears to be the human homologue of the mouse H beta 58 gene. SALT-T7 was found to have 97% identity to human 3-mercaptopyruvate sulfurtransferase and SALT-T8 was found to show homology to human interferon-inducible protein 1-8U. SALT-T9 shows approximately 90% identity to human mucin MUC 5B.

cDNA sequences encoding antigens potentially involved in small cell lung carcinoma development were isolated as follows. cDNA expression libraries were constructed with mRNA from the small cell lung carcinoma cell lines NCIH69, NCIH128 and DMS79 (all available from the American Type Culture Collection, Manassas, VA) employing the Lambda ZAP Express expression system (Stratagene, La Jolla, CA). Phagemid were rescued at random and the cDNA sequences of 27 isolated clones were determined. Comparison of the determined cDNA sequences revealed no significant homologies to the sequences of SEQ ID NO: 372 and 373. The sequences of SEQ ID NO: 364, 369, 377, 379 and 386 showed some homology to previously isolated ESTs. The sequences of the remaining 20 clones showed some homology to previously identified genes. The cDNA sequences of these clones are provided in SEQ ID NO: 363, 365-368, 370, 371, 374-376, 378, 380-385 and 387-389, wherein SEQ ID NO: 363, 366-368, 370, 375, 376, 378, 380-382, 384 and 385 are full-length sequences.

Comparison of the cDNA sequence of SEQ ID NO: 372 indicated that this clone (referred to as 128T1) is a novel member of a family of putative seven pass transmembrane proteins. Specifically, using the computer algorithm PSORT, the protein was predicted to be a type IIIA plasma membrane seven pass transmembrane protein. A genomic clone was identified in the Genbank database which contained the predicted N-terminal 58 amino acids missing from the amino acid sequence encoded by SEQ ID NO: 372. The determined full-length cDNA sequence for the 128T1 clone is provided in SEQ ID NO: 390, with the corresponding amino acid sequence being provided in SEQ ID NO: 391.

Example 7

SYNTHESIS OF POLYPEPTIDES

Polypeptides may be synthesized on a Perkin Elmer/Applied Biosystems Division 430A peptide synthesizer using FMOC chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

Example 8ISOLATION AND CHARACTERIZATION OF DNA SEQUENCES ENCODING LUNG
TUMOR ANTIGENS BY T-CELL EXPRESSION CLONING

5

Lung tumor antigens may also be identified by T cell expression cloning. One source of tumor specific T cells is from surgically excised tumors from human patients.

A non-small cell lung carcinoma was minced and enzymatically digested for
10 several hours to release tumor cells and infiltrating lymphocytes (tumor infiltrating T cells, or TILs). The cells were washed in HBSS buffer and passed over a Ficoll (100%/75%/HBSS) discontinuous gradient to separate tumor cells and lymphocytes from non-viable cells. Two bands were harvested from the interfaces; the upper band at the 75%/HBSS interface contained predominantly tumor cells, while the lower band at the
15 100%/75%/HBSS interface contained a majority of lymphocytes. The TILs were expanded in culture, either in 24-well plates with culture media supplemented with 10 ng/ml IL-7 and 100 U/ml IL-2, or alternatively, 24-well plates that have been pre-coated with the anti-CD3 monoclonal antibody OKT3. The resulting TIL cultures were analyzed by FACS to confirm that a high percentage were CD8⁺ T cells (>90% of gated population) with only a
20 small percentage of CD4⁺ cells.

In addition, non-small cell lung carcinoma cells were expanded in culture using standard techniques to establish a tumor cell line, which was later confirmed to be a lung carcinoma cell line by immunohistochemical analysis. This tumor cell line was transduced with a retroviral vector to express human CD80, and characterized by FACS
25 analysis to confirm high expression levels of CD80, class I MHC and class II MHC molecules.

The ability of the TIL lines to specifically recognize autologous lung tumor was demonstrated by cytokine release assays (IFN- γ and TNF- α) as well as ⁵¹Cr release assays. Briefly, TIL cells from day 21 cultures were co-cultured with either autologous or

allogeneic tumor cells, EBV-immortalized LCL, or control cell lines Daudi and K562, and the culture supernatant monitored by ELISA for the presence of cytokines. The TIL specifically recognized autologous tumor but not allogeneic tumor. In addition, there was no recognition of EBV-immortalized LCL or the control cell lines, indicating that the TIL lines are tumor specific and are potentially recognizing a tumor antigen presented by autologous MHC molecules.

The characterized tumor-specific TIL lines were expanded to suitable numbers for T cell expression cloning using soluble anti-CD3 antibody in culture with irradiated EBV transformed LCLs and PBL feeder cells in the presence of 20 U/ml IL-2. Clones from the expanded TIL lines were generated by standard limiting dilution techniques. Specifically, TIL cells were seeded at 0.5 cells/well in a 96-well U bottom plate and stimulated with CD-80-transduced autologous tumor cells, EBV transformed LCL, and PBL feeder cells in the presence of 50 U/ml IL-2. The specificity of these clones for autologous tumor was confirmed by ^{51}Cr microcytotoxicity and IFN- γ bioassays.

These CTL clones were demonstrated to be HLA-B/C restricted by antibody blocking experiments. A representative CTL clone was tested on a panel of allogeneic lung carcinomas and it recognized both autologous tumor and a lung squamous cell carcinoma (936T). As the only class I MHC molecule shared among these tumors was HLA-Cw1203, this indicated that this was the restriction element used by the CTL. This finding was confirmed by the recognition of a number of allogeneic lung carcinomas transduced with a retroviral vector encoding HLA-Cw1203 by the CTL.

PolyA mRNA was prepared from lung tumor LT391-06 cells using Message Maker (Life Technologies; Rockville, MD). The subsequent steps involving cDNA synthesis were performed according to Life Technologies cloning manual (SuperScript Plasmid System for cDNA Synthesis and Plasmid Cloning). Modifications to the protocol were made as follows. At the adapter addition step, EcoRI-XmnI adapters (d(AATTCGAACCCCTTCG), New England Biolabs; Beverly, MA) were substituted. Size fractionated cDNAs were ligated into the expression vector system HisMax A, B, C

(Invitrogen; Carlsbad, CA) to optimize for protein expression in all three coding frames. Library plasmids were then aliquotted at approximately 100 CFU/well into a 96-well block for overnight liquid amplification. From these cultures, glycerols stocks were made and pooled plasmid was prepared by auotmated robot (Qiagen; Valencia, CA). The
 5 concentration of the plasmid DNA in each well of the library plates was determined to be approximately 150 ng/ul. For T cell screening, approximately 80 ng of the library plasmid DNA and 80 ng of HLA-Cw1203 plasmid DNA was mixed with the lipid Fugene according the the manufacturers instructions and transfected in duplicate into COS-7 cells. After incubation at 37 C for 48 hours, the transfection mixture was removed and 10,000
 10 LT391-06 CTL were added to each well in fresh media containing human serum.

The ability of the T cells to recognize an antigen in the library was assessed by cytokine release after 6 hours (TNF-alpha, WEHI bio-assay) or after 24 hours (IFN-gamma, ELISA). Approximately $\sim 2.0 \times 10^5$ clones (in plasmid pools of 100) have been screened using this system in COS-7 cells. Three plasmid pools were identified (14F10,
 15 19A4, and 20E10) that were recognized by LT391-06 CTL. Transfection of these plasmid pools into COS-7 cells led to production of both IFN-gamma and TNF-alpha from the LT391-06 CTL significantly above background. Pools 14F10 and 19A4 were "broken down" into several hundred individual plasmid DNAs and retested. One plasmid (3D9) from pool 14F10 and 5 plasmids (2A6, 2E11, 2F12, 3F4, 3H8) from 19A4 pool were
 20 capable of reconstituting T cell recognition.

The sequencing of these plasmids identified a 7.8 kB cDNA insert (clone 14F10) and also a 2.2 kB cDNA insert (clone 19A4; SEQ ID NO:440). Clone 19A4 is contained within the 5' region of clone 14F10. BLAST search analysis against the GenBank database identified both of these sequences as having significant homology with
 25 a truncated human cystine/glutamate transporter gene. Unlike the published sequence, however, clones 14F10 and 19A4 contained a unique 5' terminus consisting of 181 nucleotides. This novel sequence replaces the published 5' region and results in the removal of the reported initiating methionine (start codon) and an additional two amino acids of the reported transporter protein. Therefore, the translated product of clones 14F10 and 19A4 is

different than the cystine/glutamate transporter protein. Furthermore, T cell recognition of other lung tumors demonstrates that this antigen is expressed by other tumors as well.

From the foregoing it will be appreciated that, although specific
5 embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

What is claimed:

1. An isolated polypeptide, comprising at least an immunogenic portion of a lung tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) sequences recited in SEQ ID NOs: 218-222, 224-226, 249, 250, 253, 256, 266, 276, 277, 282, 285, 293, 295, 298, 299, 301, 304, 306, 316, 321, 326, 333, 336, 337, 342, 353, 359, 361, 364, 369, 372, 373, 377, 379, 386, 390, 392 and 440;

(b) sequences that hybridize to a sequence recited in any one of SEQ ID NOs: 218-222, 224-226, 249, 250, 253, 256, 266, 276, 277, 282, 285, 293, 295, 298, 299, 301, 304, 306, 316, 321, 326, 333, 336, 337, 342, 353, 359, 361, 364, 369, 372, 373, 377, 379, 386, 390, 392 and 440 under moderately stringent conditions; and

(c) complements of sequences of (a) or (b).

2. An isolated polypeptide according to claim 1, wherein the polypeptide comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs: 218-222, 224-226, 249, 250, 253, 256, 266, 276, 277, 282, 285, 293, 295, 298, 299, 301, 304, 306, 316, 321, 326, 333, 336, 337, 342, 353, 359, 361, 364, 369, 372, 373, 377, 379, 386, 390, 392 and 440 or a complement of any of the foregoing polynucleotide sequences.

3. An isolated polypeptide comprising a sequence recited in any one of SEQ ID NO: 391, 393, 395, 397, 421 and 425-427.

4. An isolated polynucleotide encoding at least 15 amino acid residues of a lung tumor protein, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the tumor protein comprises an amino acid

sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NOs: 218-222, 224-226, 249, 250, 253, 256, 266, 276, 277, 282, 285, 293, 295, 298, 299, 301, 304, 306, 316, 321, 326, 333, 336, 337, 342, 353, 359, 361, 364, 369, 372, 373, 377, 379, 386, 390, 392 and 440, or a complement of any of the foregoing sequences.

5. An isolated polynucleotide encoding a lung tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NOs: 218-222, 224-226, 249, 250, 253, 256, 266, 276, 277, 282, 285, 293, 295, 298, 299, 301, 304, 306, 316, 321, 326, 333, 336, 337, 342, 353, 359, 361, 364, 369, 372, 373, 377, 379, 386, 390, 392 and 440 or a complement of any of the foregoing sequences.

6. An isolated polynucleotide, comprising a sequence recited in any one of SEQ ID NOs: 218-222, 224-226, 249, 250, 253, 256, 266, 276, 277, 282, 285, 293, 295, 298, 299, 301, 304, 306, 316, 321, 326, 333, 336, 337, 342, 353, 359, 361, 364, 369, 372, 373, 377, 379, 386, 390, 392 and 440.

7. An isolated polynucleotide, comprising a sequence that hybridizes to a sequence recited in any one of SEQ ID NOs: 218-222, 224-226, 249, 250, 253, 256, 266, 276, 277, 282, 285, 293, 295, 298, 299, 301, 304, 306, 316, 321, 326, 333, 336, 337, 342, 353, 359, 361, 364, 369, 372, 373, 377, 379, 386, 390, 392 and 440 under moderately stringent conditions.

8. An isolated polynucleotide complementary to a polynucleotide according to any one of claims 4-7.

9. An expression vector, comprising a polynucleotide according to any one of claims 4-8.

- 131

- (c) an antibody according to claim 11;
- (d) a fusion protein according to claim 12; and
- (e) a polynucleotide according to claim 16.

18. An immunogenic composition comprising an immunostimulant and at least one component selected from the group consisting of:

- (a) a polypeptide according to claim 1;
- (b) a polynucleotide according to claim 4;
- (c) an antibody according to claim 11;
- (d) a fusion protein according to claim 12; and
- (e) a polynucleotide according to claim 16.

19. An immunogenic composition according to claim 18, wherein the immunostimulant is an adjuvant.

20. An immunogenic composition according to claim 18, wherein the immunostimulant induces a predominantly Type I response.

21. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a pharmaceutical composition according to claim 17.

22. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of an immunogenic composition according to claim 18.

23. A pharmaceutical composition comprising an antigen-presenting cell that expresses a polypeptide according to claim 1, in combination with a pharmaceutically acceptable carrier or excipient.

24. A pharmaceutical composition according to claim 23, wherein the antigen presenting cell is a dendritic cell or a macrophage.

25. An immunogenic composition comprising an antigen-presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a lung tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) sequences recited in SEQ ID NOs: 217-390, 392, 394, 396, 398-420, 422-424, 428-433 and 440;

(b) sequences that hybridize to a sequence recited in any one of SEQ ID NOs: 217-390, 392, 394, 396, 398-420, 422-424, 428-433 and 440 under moderately stringent conditions; and

(c) complements of sequences of (i) or (ii);
in combination with an immunostimulant.

26. An immunogenic composition according to claim 25, wherein the immunostimulant is an adjuvant.

27. An immunogenic composition according to claim 25, wherein the immunostimulant induces a predominantly Type I response.

28. An immunogenic composition according to claim 25, wherein the antigen-presenting cell is a dendritic cell.

29. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of an antigen-presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a lung tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) sequences recited in SEQ ID NOs: 217-390, 392, 394, 396, 398-420, 422-424, 428-433 and 440;

(b) sequences that hybridize to a sequence recited in any one of SEQ ID NOs: 217-390, 392, 394, 396, 398-420, 422-424, 428-433 and 440 under moderately stringent conditions; and

(c) complements of sequences of (i) or (ii);

and thereby inhibiting the development of a cancer in the patient.

30. A method according to claim 29, wherein the antigen-presenting cell is a dendritic cell.

31. A method according to any one of claims 21, 22 and 29, wherein the cancer is lung cancer.

32. A method for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a lung tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 217-390, 392, 394, 396, 398-420, 422-424, 428-433 and 440; and

(ii) complements of the foregoing polynucleotides;

wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the antigen from the sample.

33. A method according to claim 32, wherein the biological sample is blood or a fraction thereof.

34. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated according to the method of claim 32.

35. A method for stimulating and/or expanding T cells specific for a lung tumor protein, comprising contacting T cells with at least one component selected from the group consisting of:

(a) polypeptides comprising at least an immunogenic portion of a lung tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) sequences recited in SEQ ID NOs: 217-390, 392, 396, 398-420, 422-424, 428-433 and 440;

(ii) sequences that hybridize to a sequence recited in any one of SEQ ID NOs: 217-390, 392, 396, 398-420, 422-424, 428-433 and 440 under moderately stringent conditions; and

(iii) complements of sequences of (i) or (ii);

(b) polynucleotides encoding a polypeptide of (a); and

(c) antigen presenting cells that express a polypeptide of (a);

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

36. An isolated T cell population, comprising T cells prepared according to the method of claim 35.

37. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with a binding agent that binds to a lung tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs: 217-

390, 392, 394, 396, 398-420, 422-424, 428-433 and 440 or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent; and

(c) comparing the amount of polypeptide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

38. A method according to claim 37, wherein the binding agent is an antibody.

39. A method according to claim 38, wherein the antibody is a monoclonal antibody.

COMPOSITIONS AND METHODS FOR THE THERAPY
AND DIAGNOSIS OF LUNG CANCER

ABSTRACT OF THE DISCLOSURE

Compositions and methods for the therapy and diagnosis of cancer, such as lung cancer, are disclosed. Compositions may comprise one or more lung tumor proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a lung tumor protein, or a T cell that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of diseases such as lung cancer. Diagnostic methods based on detecting a lung tumor protein, or mRNA encoding such a protein, in a sample are also provided.

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants : Steven G. Reed et al.
Filed : September 20, 2000
For : COMPOSITIONS AND METHODS FOR THE THERAPY AND
DIAGNOSIS OF LUNG CANCER

Docket No. : 210121.475C7

Date : September 20, 2000

Box Patent Application
Assistant Commissioner for Patents
Washington, D.C. 20231

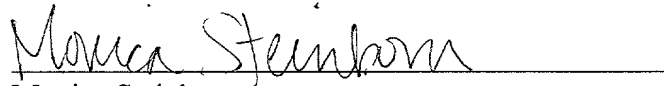
DECLARATION

Sir:

I, Monica Steinborn, in accordance with 37 C.F.R. § 1.821(f) do hereby declare that, to the best of my knowledge, the content of the paper entitled "Sequence Listing" and the computer readable copy contained within the floppy disk are the same.

I declare further that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true and further that these statements are made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Dated this 20th day of September, 2000.



Monica Steinborn
Biotechnology Paralegal

701 Fifth Avenue, Suite 6300
Seattle, WA 98104-7092
(206) 622-4900
FAX (206) 682-6031

000250 "0272960

<110> Reed, Steven G.	Henderson, Robert A.
Lodes, Michael J.	Fling, Steven P.
Mohamath, Raodoh	Algate, Paul A.
Secrist, Heather	Indirias, Carol Yoseph
Benson, Darin R.	

<130> 210121.475C7

<141> 2000-09-20

<160> 440

<170> FastSEQ for Windows Version 3.0

<211> 339

<213> Homo sapien

gtactcagac	aggatagtc	tcatgtagca	caaagcamat	cctgtttcta	tacttgtagt	60
ttgctctcac	tcagtggcat	ratcattact	atacagtgt	gaatgttrtt	atgtagcata	120
gatgtggggt	ctctagccca	cagctctsta	cctttgtcta	gcactcctgt	cctcatacct	180
ragtggcctg	tccatcagca	tgtttctcat	ctactttgct	tgtccagtc	actgtgggtcc	240
tcccttgccc	tctcccttat	gtggcagagt	ggaaccagct	gtcctgagac	ttgagttcaa	300
catctggttc	qcccatytoc	atgtttatga	tctgagtac			339

<211> 698

<213> Homo sapien

<221> misc feature

<222> (1) ... (698)

<223> n = A, T, C or G

gtactcagac	cacgactgca	ttttctccac	tgctgacggg	tctaatacca	gctgcttccc	60
tttcttgagg	gcagagctng	tgaccttgag	aaagtgacct	gtgaccatca	tgtgggtagt	120
gagctgctgc	aaggtgtcat	gggagctccc	acactccatg	cactttwaga	tctgggactt	180
gcaggcctca	ractgccagg	tgtagctcgc	tccatttttg	tagccatagc	gsttggttga	240
ggacaactgc	aagttggcgt	tcttctgaga	agaaaaagaa	tctgcaaaag	atcctgtggt	300
tgaatcgggg	gaacacggcc	gattgacatc	aaaaacgcgt	ttcttagccc	gggtgaccat	360
tttcgaggaa	atggttgggg	actggctcct	tcaaaggcac	tttttggtta	tgttttgttt	420
yaatcatgtk	gacgctccaa	tcttggragg	gaatcgaang	rantcncnc	caaaacatrc	480
stttcagraa	ccttttgarc	atcctctttt	ttccgtrtcc	cgqmaargcc	cytttccckg	540

```

ggctttgaaa wyagcctsgt tgggttctta aattaccart ccacnwggtg gaattccccc 600
ggccccctgc ccgggtccaa ccaatttttg graaaacccc cncansccgt tkggantgen 660
acaacntggn ntttttcntt tegtntccc ctngaacc 698

```

```

<210> 3
<211> 697
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(697)
<223> n = A,T,C or G

```

```

<400> 3
gtactcagac ccccaacctc gaacagccag aagacaggtt gtctcctggg ccttggacac 60
agccngccag gccattgaag ganaagcaaa gacgaagcga accatctctc tccattgtgg 120
gggccaagta gctgcantan ccttcagtcc cagttgcatt gggttaaaga gctcatacat 180
actatgtgtn aggggtacag aagcttttcc tcatagggca tgagctctcc nagagttgac 240
cttttgccn aaacttgggt ttctgtggtt cataaagttt ggatatgtat tttttttcaa 300
atggaanaaa atccgtatth ggcaaaaaga ctccaggggg atgatactgt ccttgccact 360
tacagtccaa angatnttcc ccaaagaata gacatttttt cctctcatca cttctggatg 420
caaaatcttt ttttttttcc ctttctcgca ccnccccaga ccccttnnag gttnaaccgc 480
ttcccatctc ccccatccca cagcatnttg aattngcann ncgttgntgg tcgggtcccn 540
nccgaaaggg tntttttatt cggggtnctg anttnnaac cnctnagttg aatccgcggg 600
gcggccnngn ggggttnnacc atgntgggga naactncccn ccgcgnttgg aatgccanag 660
cctgaaant tttcttttgg tcgcccccn gagattc 697

```

```

<210> 4
<211> 712
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(712)
<223> n = A,T,C or G

```

```

<400> 4
gtactcagac aaccaatagg tgtgttyctc anactgaaa cacaaaaaga ttctagctna 60
taatgttsaa tgggttgaggg tttaagtgat cttggtatgt tngatttagc agcgatnggc 120
cgggtgcggt ggctcacgca tgtatcccag cactttggga ggccgaggca ggaggatcac 180
ctgaggtcag gagtttgaga ccagcctggc cgacatggtn aaaccccgtc tctactanga 240
atacanaaat tagcccgggc atagtggcgc gtgcctrtga cctcsgctac tttggggatt 300
ctcctgagga agaattgctt gaactcaggg aagtggargt ttgcagtgag cttaaatcaa 360
gccactggca ctcccagcct gggktaacag agccamgact ctkgccgaaa aaaaaaama 420
cgacggagaa nmagntctgt tattccatgg gaaattkgaa tttccttctt tkaaatatct 480
taaaatnggt cctcctwaaa aaagttcggc tggggcccgc tggctcacat tttkttaycc 540
cycccccttt tggggarggc caarggccg kttgawtnnc ccttgagggg ccanaactcc 600
agnaaccrgn ccggggccar smgwkgtstr armccctttc cyyccmaraa aawwcsmaa 660
wwttyccsc cygsykggct ggkasckgtt myyyyygntm csyagcttgc tt 712

```

```

<210> 5
<211> 679

```

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(679)
<223> n = A,T,C or G

<400> 5

gtactcagac	cacctcacat	gcagggtgag	aaacatggag	tgtgcggcag	catcctcctc	60
acatcccttt	gtgagcacgg	ctgctccgga	atactgacca	tctgggctag	cacgacctaa	120
cagagggttc	tgcaggatgt	gctattttaa	agcagctggg	tgcaacttgt	gaaaacggga	180
atctngaagc	agaacatgtn	atcagcgatg	gctgggattg	gtggacagga	ttgacaggag	240
tatttgaggc	tctaccaggc	ctgtctacag	gacagcttca	tcgaagggac	atTTTTtaac	300
ctgttatTTT	anatnccaca	tatntTTTT	aatgctnaag	catacagggt	gaatttcttg	360
atcgtaacta	ctagtgactt	ctgagggtta	cagttngaag	atgttctcnn	aggtttatca	420
agttntgtta	ttgatgatng	gtaatctaca	cctctggaag	ctgtngaag	tgaaaaagat	480
nctntcanct	gaccagtttg	nagggcactc	tcttctggna	agnaatccgn	ccaaaaaat	540
tgtttcnagg	gggcntgggg	ggtttaaaaa	aatgtttctn	ttncntaaa	aatgtttacc	600
cnnctattga	aaaaatgggg	gtcngggggg	gcttnaaatc	cccnanttnt	gaatnttnta	660
tccggaanct	tggtttccc					679

<210> 6
<211> 369
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(369)
<223> n = A,T,C or G

<400> 6

tcagtccagt	catgggtcct	ataagagaag	tcactctgtg	agtttccatg	gaggaagaaa	60
aagcttcatt	tctttaccct	gcagcaacag	cggagggagg	gagagcctat	cttctttgca	120
aattcattaa	ctttgtggtt	gaaggggagca	gcgtcngaaa	ctgcttttagc	acagtgggag	180
gaaaacaaac	agattcatct	ccggaaacca	aaggaaaggg	tragtggggt	tttattagcc	240
agctgtatcc	tagatgggtca	atttccagtg	gatgaataca	ccttacgtac	gtttctcttg	300
cttctacct	nggcctgatc	agctnggcac	ttraatcatt	ccgtnggggt	wgctgtnaca	360
ctggactga						369

<210> 7
<211> 264
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(264)
<223> n = A,T,C or G

<400> 7

tgctggatra	gggatggggc	acggggagcac	agatmgactt	taactgcccc	cacgttntcm	60
aggaaaggat	tacaggcggtg	agccactgcg	ccggcctct	tctccacttt	cataggttcc	120

agtctctggt tcttctttct cagtttggtg tttttgcttc ttaaamtatg gagatnagaa 180
 tgaacactac actcggaaac aggaagccct gcctggcgcc tctgtcacct gtctaggggc 240
 ttcttctcac tgagtcaccc agca 264

<210> 8
 <211> 280
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(280)
 <223> n = A,T,C or G

<400> 8
 acctcaactg ccanaacan aactgttgta caagatttga ggatttaaca atatttcaca 60
 tgaaatattt cagacctacg ngagggctta aagacnaatt aaatgagcac cngtgtgccc 120
 accgcccna ttaagaatta gagcaagcag tgagggtgaag ccttgctcctt gcttttaaca 180
 tagaaagtga tccaaattca ccaaacttga cttnnggttt tgcagtgtgg cctcctgatt 240
 ctgacnctg gcgaaacatt tgatgggcaa aaaaaaaaaa 280

<210> 9
 <211> 449
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(449)
 <223> n = A,T,C or G

<400> 9
 tcgtcaactc caggatggct ttgaaaatna atggacacag atctctcctg ttttgatrat 60
 ntgcagtgt natgactggc tttgcagtn attttgattc aggcaacaga tggttcctttt 120
 ggttcctgt ctcocatggg cgtcatttca tgttgctcctc tgccttcccc cagatattct 180
 aagttcagga cacaagcttc tggcccatgc agagcagagg ccatgagggg tcacagcatg 240
 ggtacgggag gaaacactgg gctnaccag atnctggact tgagtcttgc ctctgctgct 300
 tgctgcacag cttctgtcat ggtgctaaac ctgtgacctg cctcacaggc ttagagcatg 360
 cccgtagaag tactctnaac taaratgctt tccacaaatg agatggtttc atgaaaactt 420
 caaatagagg gcctgggcaa aaaaaaaaaa 449

<210> 10
 <211> 538
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(538)
 <223> n = A,T,C or G

<400> 10
 tttttttttt ttcccaaagg cctcaraaca ctagtcttct aattccaagc agaaagttac 60
 atccgcccgg atacatgcc cttgggttga taaatcaaaa tacagcatcc ttcagatccc 120

09667170 "096000

```
<210> 11
<211> 543
<212> DNA
<213> Homo sapien
```

<400> 11

```
<210> 12
<211> 329
<212> DNA
<213> Homo sapien
```

<400> 12

```
<210> 13
<211> 314
<212> DNA
<213> Homo sapien
```

<400> 13

```
<210> 14
<211> 691
<212> DNA
<213> Homo sapien
```

```
<220>  
<221> misc_feature  
<222> (1)...(691)  
<223> n = A,T,C or G
```

```
<210> 15
<211> 355
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(355)
<223> n = A,T,C or G
```

```
<210> 16
<211> 522
<212> DNA
<213> Homo sapien
```

<220>
 <221> misc_feature
 <222> (1)...(522)
 <223> n = A,T,C or G

<400> 16
 tcagtccagt gaggtggaag acttcgaggc tcgtgggagc cgcttctcca agtctgctga 60
 tgagagacag cgcagtctgg tgcagcgtan ggacgaactc ctccagcaag ctgcgagacg 120
 tttcttgaac aaaagttctg aagatgatgc ggccctcagag agcttcctcc cctcggaagg 180
 tgcgtcctct gaccccgatga cctnccgtcg aangatgctg gctgccgccg cggaacggan 240
 gcttcagaag cagcagacct cctnccgtc ccttgccctc ctccagctgcc tcctgcccgc 300
 tgtgcccggc tgactggagg aggcctgtcc aattctgccc gcccctatgga aaagcgggct 360
 tgactgcatt gccgctgtat naaagcatgt ggtcttacag tgttnggacn gctnatnaat 420
 ttnatcctnc tntgtaatac ttcctatgtg acatttctct tccccttggg aacactgcan 480
 attttaactg tgagtttgat ctcttctnct gttactggac tg 522

<210> 17
 <211> 317
 <212> DNA
 <213> Homo sapien

<400> 17
 gtgtcgcgaa ttcgcggtgg tgctaagaaa aggaagaaga agtcttacac cactcccaag 60
 aaggataagc accagagaaa gaagggttcag ccggccgtcc tgaaatatta taagggtggat 120
 gagaatggca aaattagttg ccttcgtcga gagtgcacct ctgatgaatg tgggtgctggg 180
 gtgtttatgg caagtcactt tgacagacat tattgtggca aatgttgtct gacctactgt 240
 ttcaactaac cagaagacaa gtaactgtat gagttaatta aagacatgaa ctaaaaaaaa 300
 aaaaaaaaaa actcgag 317

<210> 18
 <211> 392
 <212> DNA
 <213> Homo sapien

<400> 18
 tggagatttc taatgaggtg aggaagttcc gtacattgac agaattgatc ctccgatgctc 60
 aggaacatgt taaaaatcct taaaaaggca aaaaactcaa gaaacaccca gacttcccca 120
 agaagccctc gaccccttat ttccgcttct tcatggagaa gcgggccaaag tatgcgaaac 180
 tccaccctca gatgagcaac ctggacctga ccaagattct gtccaagaaa tacaaggagc 240
 ttccggagaa gaagaagatg aaatatgttc cggacttcca gagaagagaa acaggagttc 300
 gagcgaaacc tggcccgatt caggagggat cccccccacc ttatccagaa tgccaagaat 360
 cggacatccc agagaagccc caagaccccc cg 392

<210> 19
 <211> 2624
 <212> DNA
 <213> Homo sapien

<400> 19
 gaaacagtga gaaggagatt cctgtgctca atgagctgcc agtcccatg gtggcccgct 60
 acattcgcatt aaacctcag tcttggtttg ataacgggag catctgcatg aggatggaga 120
 tcttgggctg cccactgccg gatcctaata actattatca ccgacgtaat gagatgacca 180
 ccacggatga cctggatttt aagcaccaca actattagga aatgcgccag ttgatgaagg 240
 ttgtcaatga aatgtgcccc aatattacca ggatttataa cattggcaaa agccaccagg 300

gcctgaaatt gtatgcggtg gagatctctg accatcctgg ggaacatgaa gttggtgagc 360
 ccgagttcca ctacatcgca ggggcccacg gcaatgaggt tctgggacga gaactgctgc 420
 tgctgctgct gcaacttctc tgccaggaat actcggcgca gaacgcacgc atcgtccgct 480
 tgggtggagga gactcgaatc cacatttctac cctccctcaa tcctgatggc tatgagaagg 540
 cctatgaagg aggttccgag ttgggagggt ggtccctggg acgttgacc catgatggca 600
 tcgatatcaa caacaacttt ccggatttaa actcgtctgt ctgggaggca gaggaccagc 660
 agaatgcccc aagggaaggtc cccaaccact acattgccat ccctgagtgg tttctgtctg 720
 agaatgccac agtggccaca gagaccagag ccgtcatcgc ctggatggag aagatcccgt 780
 ttgtgctggg aggcaaccta caggggggtg agctggtcgt ggcatacccc tatgacatgg 840
 tgcggtccct gtggaagacc caggagcaca ccccaacacc tgatgatcat gtgttccgct 900
 ggctggcgta ttctacgcc tccactcacc gcctcatgac agatgccagg aggcgagtgt 960
 gccacacgga agattttcag aaggaggagg gcaccgtcaa tggggcttcc tggcacacag 1020
 tggctggaag tctaaacgat ttcagctacc tccatacaaa ctgctttgag ctgtccatct 1080
 acgtgggctg tgataaatac ccacacgaga gcgagctgcc ggaggaatgg gagaataacc 1140
 gggagtctct gattgtgttc atggagcagg ttcactcgagg catcaaaggc atagtgaag 1200
 atttacaagg gaaagggatt tcaaattgctg tcatctctgt ggaaggtgtt aaccatgaca 1260
 tccggacagc cagcgtatggg gattactggc gtctactgaa ccctggcgaa tatgtgttca 1320
 cagccaaggc ggaaggcttt atcacttcca ccaagaactg catggttggc tatgatattg 1380
 gagctactcg gtgtgacttc accctcacia agaccaacct ggctaggata agagaaatta 1440
 tggagacatt tgggaagcag cctgtcagcc taccctccag gcgcctgaag ctgcggggac 1500
 ggaaaaggcg gcagcgtggg tgaccctgtc ggacacttga gacatacccc agaccgtgca 1560
 aataaaaatc cactccagta gtaactctgt agcaggcttt ccctgttgtt ttgactgtaa 1620
 ttcaagagac actcaggagc atacctgcat ggcttggctg accccaaagg ggagggtcgg 1680
 tggctcaggg tgttttgttt tttgtttttt gttttttcct ttgttctcat ttatccaaat 1740
 acctgaaca gagcagcaga gaaaggccgg tggcagttag ggaattaatt cagtgaagta 1800
 gtctgagatt ctaaaaaggg tgcttgacca ctggccagga agggaaatca ggcttcccc 1860
 catttgctg acattcaagg ttcccagtg atttgcaagt ggcacagtgt acattgcagc 1920
 acccagggaa tcctttgccc cagatgttat catttgagat gctcttatgc agcctaagaa 1980
 aatccatcct ctctggcccc aggggacaag ccaagctgct atgtacacac tcggtgttct 2040
 attgacaata gaggcattta ttaccaagtg tgcactcgtg agtcctaaat cagctctgtt 2100
 cctttttcca acaaaagctt tcttcttaag agcagacaga agtggagagc acccaagaat 2160
 gagtgcctgg cagcagaccc tgggggaggg ggcttgctat ccagaaaagc ccctaaaccc 2220
 tttgctgctc cattagccct ggggtgagga gagccagaca tgtaggagg ccagagcagt 2280
 cagtcagggc atcttgaaa agaccttgaa ggaagcaaac cctgggttcc tttgctcca 2340
 gaatgtgaga gctccaagtt ggccccaatc agggggggag taatgatgaa catacagacg 2400
 gccacatctt gccaatcaag catcatctga tgaaaaagaa agcaatctta ggattacctg 2460
 ggacacgtca gtctgggaga ggtggttgaa tcattgtgta agggaatagt gtatctaata 2520
 tgtgttgatc ctgctgcctt gttgacctgg agagaatgaa acaaacaaac acataaacaa 2580
 ataaagcaaa tggttaagatt aaaaaaaaaa aaaaaaaact cgag 2624

<210> 20

<211> 488

<212> DNA

<213> Homo sapien

<400> 20

ctttcaaccc gcgtcgcgg gctccagccc cgcgcgcccc cacccttgc cctccggcg 60
 gctccgcagg gtgaggtggc ttgacccccg ggttgccccg ccagcacgac cgaggaggtg 120
 gctggacagc tggaggatga acggagaagc cgactgcccc acagacctgg aaatggccgc 180
 cccagaggc caagaccgtt ggtcccagga agacatgctg actttgctgg aatgcatgaa 240
 gaacaacctt ccatccaatg acagctccca gttcaaaacc acccaaacac acatggaccg 300
 ggaaaaagtt gcattgaaag acttttctgg agacatgtgc aagctcaaata gggtcgagat 360
 ctctaatagag gtgaggaagt tccgtacatt gacagaattg atcctcgata ctcaggaaca 420
 tgtttaaaaat ccttacaagg gcaaaaaatc aagaaacacc ccgacttccc cgagaaagcc 480

cctaaccc

488

<210> 21
 <211> 391
 <212> DNA
 <213> Homo sapien

<400> 21

atggaattgt	ggttttctct	ttgggatcaa	tggtctcaga	aattccagag	aagaaagctg	60
tggcgattgc	tgatgctttg	ggcaaaatcc	ctcagacagt	cctgtggcgg	tacactggaa	120
cccgaccatc	gaatcttgcg	aacaacaaga	tacttggtca	gtggctaccc	caaaacgac	180
tgcttggtca	cccaatgacc	cgtgccttta	tcacccatgc	tagttcccat	ggtgttaatg	240
aaagcatatg	caatggcggt	cccatgggtga	tgataccctt	atttgggtgat	cagatggaca	300
atgcaaagcg	caggagagct	aaggagagctg	gagtgaccct	gaatgttctg	gagatgactt	360
ctgaagatct	agaagatgct	ctgaagagca	g			391

<210> 22
 <211> 1320
 <212> DNA
 <213> Homo sapien

<400> 22

aatctgctgg	gaatttcttg	ggttgacagc	tcttgatcc	ctattttgaa	cagtggtagt	60
gtcctggatt	acttttcaga	aagaagtaat	cctttttatg	acagaacatg	taataatgaa	120
gtggtcaaaa	tgagagggct	aacattagaa	cacttgaatc	agatgggttg	aatcgagtac	180
atccttttgc	atgctcaaga	gccattctt	ttcatcttc	ggaagcaaca	gcggcagtc	240
cctgccaag	ttatccact	agctgattac	tatatcattg	ctggagtgat	ctatcaggca	300
ccagacttgg	gatcagttat	aaactctaga	gtgcttactg	cagtgcagtg	tattcagtca	360
gcttttgatg	aagctatgtc	atactgtcga	tatcatcctt	ccaaagggtg	ttggtggcac	420
ttcaaagatc	atgaagagca	agataaagtc	agacctaaag	ccaaaaggaa	agaagaacca	480
agctctattt	ttcagagaca	acgtgtggat	gctttacttt	tagacctcag	acaaaaat	540
ccacccaaat	ttgtgcagct	aaagcctgga	gaaaagcctg	ttccagtgga	tcaaacaaag	600
aaagaggcag	aacctatacc	agaaactgta	aaacctgagg	agaaggagac	cacaaagaat	660
gtacaacaga	cagtgaagtgc	taaaggcccc	cctgaaaaac	ggatgagact	tcagtgaagta	720
ctggacaaaa	gagaagcctg	gaagactcct	catgctagtt	atcatacctc	agtactgtgg	780
ctcttgagct	ttgaagtact	ttattgtaac	cttcttattt	gtatggaatg	cgcttatttt	840
ttgaaaggat	attaggccgg	atgtgggtggc	tcacgcctgt	aatcccagca	ctttgggagg	900
ccatggcggg	tggatcactt	gaggtcagaa	gttcaagacc	agcctgacca	atatgggtgaa	960
accccgctctc	tactaaaaat	acaaaaatta	gccgggcgtg	gtggcgggcg	cccatagtcc	1020
cagctactcg	ggaggctgag	acaggagact	tgcttgaacc	cgggaggtgg	aggttgccct	1080
gagctgatca	tcctgctgtt	gcactccagc	ttgggcgaaa	gagcgagact	ttgtctctat	1140
aaagaaggaa	agatattatt	cccatcatga	tttcttgtga	atatttgtaa	tatgtttttt	1200
gtaacctttc	ctttcccgga	cttgagcaac	ctacacactc	acatgtttta	tggtagatat	1260
gttttaaaagc	aagataaagg	tattgggtttt	aaaaaaaaaa	aaaaaaaaaa	aaaactcgag	1320

<210> 23
 <211> 633
 <212> DNA
 <213> Homo sapien

<400> 23

ctaagggcag	tgaaggtgaa	aaccctctca	cgggtcccagg	gagggagaag	gaaggcatgc	60
tgatgggggt	taagccgggg	gaggacgcac	cggggcctgc	tgaagacctt	gtgagaagat	120
ctgagaaaaga	tactgcagct	gttgtctcca	gacagggcag	ctccctgaac	ctctttgaag	180

0956170-095000

atgtgcagat	cacagaacca	gaagctgagc	cagagtccaa	gtctgaaccg	agacctccaa	240
tttcctctcc	gagggctccc	cagaccagag	ctgtcaagcc	ccgacttcat	cctgtgaagc	300
caatgaatgc	cacggccacc	aagggttgcta	actgcagctt	gggaactgcc	accatcatcg	360
gtgagaactt	gaacaatgag	gtcatgatga	agaaatacag	ccccctggac	cctgcatttg	420
catatgcgca	gctgacccac	gatgagctga	ttcagctggg	cctcaaacag	aaggaaacga	480
taagcaagaa	ggagttccag	gtccgcgagc	tggaagacta	cattgacaac	ctgctcgtca	540
gggtcatgga	agaaaccccc	aatatcctcc	gcaccccgac	tcaggttggc	aaaaaagcag	600
gaaagatgta	aattagcaga	aaaaaaactc	gag			633

<210> 24

<211> 1328

<212> DNA

<213> Homo sapien

<400> 24

gtaaacgctc	tcggaattat	ggcggcggtg	gatatccgag	acaatctgct	gggaattttct	60
tgggttgaca	gctcttgat	ccctattttg	aacagtggta	gtgtcctgga	ttacttttca	120
gaaagaagta	atccttttta	tgacagaaca	tgtaataatg	aagtgggtcaa	aatgcagagg	180
ctaacattag	aacacttgaa	tcagatgggt	ggaatcgagt	acatcctttt	gcatgctcaa	240
gagcccattc	ttttcatcat	tcggaagcaa	cagcggcagt	ccccgccc	agttatccca	300
ctagctgatt	actatatcat	tgctggagtg	atctatcagg	caccagactt	gggatcagtt	360
ataaactcta	gagtgcctac	tgacagtgc	ggtattcagt	cagcttttga	tgaagctatg	420
tcatactgtc	gatatcatcc	ttccaaagg	tattgggtggc	acttcaaaga	tcataagag	480
caagataaag	tcagaccta	agccaaaagg	aaagaagaac	caagctctat	ttttcagaga	540
caacgtgtgg	atgctttact	tttagacctc	agacaaaaaa	tttccaccca	aatttgtgca	600
gtggatcaaa	caaagaaaaga	ggcagaacct	ataccagaaa	ctgtaaaacc	tgaggagaag	660
gagaccacaa	agaatgtaca	acagacagtg	agtgtctaaag	gccccctga	aaaacggatg	720
agacttcagt	gagtactgga	caaaagagaa	gcctggaaga	ctcctcatgc	tagttatcat	780
acctcagtac	tgtggctctt	gagctttgaa	gtactttatt	gtaaccttct	tatttgtatg	840
gaatgcgctt	atttttttga	aaggatatta	ggcgggatgt	ggtggctcac	gcctgtaatc	900
ccagcacttt	gggaggccat	ggcgggtgga	tcacttgagg	tcagaagttc	aagaccagcc	960
tgaccaatat	ggtgaaaccc	cgtctctact	aaaaatacaa	aaattagccg	ggcgtggtgg	1020
cgggcgccc	tagtcccagc	tactcgggag	gctgagacag	gagacttgct	tgaacccggg	1080
aggtggaggt	tgccctgagc	tgattatcat	gctgttgac	tccagcttgg	gcgacagagc	1140
gagactttgt	ctcaaaaaag	aagaaaagat	attattccca	tcattgatttc	ttgtgaatat	1200
ttgtgatatg	tcttctgtaa	cctttcctct	cccggacttg	agcaacctac	acactcacat	1260
gtttactggt	agatatgttt	aaaagcaaaa	taaaggtatt	tgtataaaaa	aaaaaaaaaa	1320
aaactcga						1328

<210> 25

<211> 1758

<212> DNA

<213> Homo sapien

<400> 25

gttttttttt	tttttttttt	aaagagttgc	aacaattcat	ctttatttct	tatttttcctc	60
tggagatgca	gaatttggtg	tattttcacc	caagtatatt	tgggatagtt	ggctcctcgc	120
tgggtcagga	tggttggtg	ccttctcccc	tggcatggtt	ctcttctctg	cagggcgagg	180
ggcagggagc	tagtaaaacc	tcgcaatgac	agccgcaatg	gcagacccaa	tgagagccag	240
gatgaacttg	gtcaatccgg	agagtccagt	tgctcccagt	gactgcagag	tagccacaag	300
gctgcccag	gcaactccac	ccccattggc	aatggccgcc	gcggacatca	tcttggtgc	360
tatggaggac	gagggcgattc	ccgcgcgagt	gaagcccatg	gcactgagtg	gcggcggtgg	420
atatccgaga	caatctgctg	ggaattttct	gggttgacag	ctcttggtac	cctatttttg	480
acagtggtag	tgtcctggat	tactttttcag	aaagaagtaa	tcctttttat	gacagaacat	540

```
<210> 26
<211> 493
<212> DNA
<213> Homo sapien
```

```
<210> 27
<211> 1331
<212> DNA
<213> Homo sapien
```

<400> 27						
ggtggatatt	cgagacaatc	tgtctgggaat	ttcttggggtt	gacagctctt	ggatccctat	60
tttgaacagt	ggtagtgtcc	tggattactt	ttcagaaaga	agtaatcctt	tttatgacag	120
aacatgtaat	aatgaagtgg	tcaaaatgca	gaggctaaca	ttagaacact	tgaatcagat	180
ggttggaaatc	gagtacatcc	ttttgcatgc	tcaagagccc	attcttttca	tcattcgga	240
gcaacagcgg	cagtcccttg	cccaagttat	cccactagct	gattactata	tcattgctgg	300
agtgatctat	caggcaccag	acttgggatc	agttataaac	tctagagtgc	ttactgcagt	360
gcatggtatt	cagtcagctt	ttgatgaagc	tatgtcatac	tgtcgatatc	atccttccaa	420
agggatttgg	tggcacttca	aagatcatga	agagcaagat	aaagtcaagc	ctaaagccaa	480
aaggaaagaa	gaaccaagct	ctatttttca	gagacaagct	gtggatgctt	tactttttaga	540
cctcagacaa	aaatttccac	ccaaatttqt	gcagctaaag	cctggagaaa	agcctgttcc	600

```
<210> 28
<211> 1333
<212> DNA
<213> Homo sapien
```

```
<210> 29
<211> 813
<212> DNA
<213> Homo sapien
```

<400> 29						
ctgagctgca	cttcagcgaa	ttcacctcgg	ctgtggctga	catgaagaac	tccgtggcgg	60
accgagacaa	cagccccagc	tcctgtgctg	gcctcttcat	tgcttcacac	atcggggttg	120
actggcccg	ggtctgggtc	caacctggaca	tcgctgctcc	agtgcattgt	ggcgagcggag	180
ccacaggttt	tgggtgggct	ctcctactgg	ctcttttttg	ccgtgcctcc	gaggaccgcg	240

115				120				125							
Pro	Pro														
130															
<210>	34														
<211>	506														
<212>	PRT														
<213>	Homo sapien														
<400>	34														
Asn	Ser	Glu	Lys	Glu	Ile	Pro	Val	Leu	Asn	Glu	Leu	Pro	Val	Pro	Met
1				5					10					15	
Val	Ala	Arg	Tyr	Ile	Arg	Ile	Asn	Pro	Gln	Ser	Trp	Phe	Asp	Asn	Gly
			20					25					30		
Ser	Ile	Cys	Met	Arg	Met	Glu	Ile	Leu	Gly	Cys	Pro	Leu	Pro	Asp	Pro
		35					40					45			
Asn	Asn	Tyr	Tyr	His	Arg	Arg	Asn	Glu	Met	Thr	Thr	Thr	Asp	Asp	Leu
	50					55					60				
Asp	Phe	Lys	His	His	Asn	Tyr	Lys	Glu	Met	Arg	Gln	Leu	Met	Lys	Val
65					70					75				80	
Val	Asn	Glu	Met	Cys	Pro	Asn	Ile	Thr	Arg	Ile	Tyr	Asn	Ile	Gly	Lys
				85					90					95	
Ser	His	Gln	Gly	Leu	Lys	Leu	Tyr	Ala	Val	Glu	Ile	Ser	Asp	His	Pro
			100					105					110		
Gly	Glu	His	Glu	Val	Gly	Glu	Pro	Glu	Phe	His	Tyr	Ile	Ala	Gly	Ala
		115					120					125			
His	Gly	Asn	Glu	Val	Leu	Gly	Arg	Glu	Leu	Leu	Leu	Leu	Leu	Leu	His
	130					135					140				
Phe	Leu	Cys	Gln	Glu	Tyr	Ser	Ala	Gln	Asn	Ala	Arg	Ile	Val	Arg	Leu
145					150				155						160
Val	Glu	Glu	Thr	Arg	Ile	His	Ile	Leu	Pro	Ser	Leu	Asn	Pro	Asp	Gly
				165					170					175	
Tyr	Glu	Lys	Ala	Tyr	Glu	Gly	Gly	Ser	Glu	Leu	Gly	Gly	Trp	Ser	Leu
			180					185					190		
Gly	Arg	Trp	Thr	His	Asp	Gly	Ile	Asp	Ile	Asn	Asn	Asn	Phe	Pro	Asp
		195					200				205				
Leu	Asn	Ser	Leu	Leu	Trp	Glu	Ala	Glu	Asp	Gln	Gln	Asn	Ala	Pro	Arg
	210					215					220				
Lys	Val	Pro	Asn	His	Tyr	Ile	Ala	Ile	Pro	Glu	Trp	Phe	Leu	Ser	Glu
225					230					235					240
Asn	Ala	Thr	Val												

```
<210> 35
<211> 96
<212> PRT
<213> Homo sapien
```

```
<210> 36
<211> 129
<212> PRT
<213> Homo sapien
```

			<400>	36												
Gly	Ile	Val	Val	Phe	Ser	Leu	Gly	Ser	Met	Val	Ser	Glu	Ile	Pro	Glu	
1				5					10					15		
Lys	Lys	Ala	Val	Ala	Ile	Ala	Asp	Ala	Leu	Gly	Lys	Ile	Pro	Gln	Thr	
			20					25					30			
Val	Leu	Trp	Arg	Tyr	Thr	Gly	Thr	Arg	Pro	Ser	Asn	Leu	Ala	Asn	Asn	
		35					40					45				
Thr	Ile	Leu	Val	Gln	Trp	Leu	Pro	Gln	Asn	Asp	Leu	Leu	Gly	His	Pro	
	50					55					60					


```
<210> 37
<211> 238
<212> PRT
<213> Homo sapien
```

```
<210> 38
<211> 202
<212> PRT
<213> Homo sapien
```

```
<210> 38
<211> 202
<212> PRT
<213> Homo sapien
```

Lys Gly Ser Glu Gly Glu Asn Pro Leu Thr Val Pro Gly Arg Glu Lys
 1 5 10 15
 Glu Gly Met Leu Met Gly Val Lys Pro Gly Glu Asp Ala Ser Gly Pro
 20 25 30
 Ala Glu Asp Leu Val Arg Arg Ser Glu Lys Asp Thr Ala Ala Val Val
 35 40 45
 Ser Arg Gln Gly Ser Ser Leu Asn Leu Phe Glu Asp Val Gln Ile Thr
 50 55 60
 Glu Pro Glu Ala Glu Pro Glu Ser Lys Ser Glu Pro Arg Pro Pro Ile
 65 70 75 80
 Ser Ser Pro Arg Ala Pro Gln Thr Arg Ala Val Lys Pro Arg Leu His
 85 90 95
 Pro Val Lys Pro Met Asn Ala Thr Ala Thr Lys Val Ala Asn Cys Ser
 100 105 110
 Leu Gly Thr Ala Thr Ile Ile Gly Glu Asn Leu Asn Asn Glu Val Met
 115 120 125
 Met Lys Lys Tyr Ser Pro Ser Asp Pro Ala Phe Ala Tyr Ala Gln Leu
 130 135 140
 Thr His Asp Glu Leu Ile Gln Leu Val Leu Lys Gln Lys Glu Thr Ile
 145 150 155 160
 Ser Lys Lys Glu Phe Gln Val Arg Glu Leu Glu Asp Tyr Ile Asp Asn
 165 170 175
 Leu Leu Val Arg Val Met Glu Glu Thr Pro Asn Ile Leu Arg Ile Pro
 180 185 190
 Thr Gln Val Gly Lys Lys Ala Gly Lys Met
 195 200

<210> 39

<211> 243

<212> PRT

<213> Homo sapien

<400> 39

Val Asn Ala Leu Gly Ile Met Ala Ala Val Asp Ile Arg Asp Asn Leu
 1 5 10 15
 Leu Gly Ile Ser Trp Val Asp Ser Ser Trp Ile Pro Ile Leu Asn Ser
 20 25 30
 Gly Ser Val Leu Asp Tyr Phe Ser Glu Arg Ser Asn Pro Phe Tyr Asp
 35 40 45
 Arg Thr Cys Asn Asn Glu Val Val Lys Met Gln Arg Leu Thr Leu Glu
 50 55 60
 His Leu Asn Gln Met Val Gly Ile Glu Tyr Ile Leu Leu His Ala Gln
 65 70 75 80
 Glu Pro Ile Leu Phe Ile Ile Arg Lys Gln Gln Arg Gln Ser Pro Ala
 85 90 95
 Gln Val Ile Pro Leu Ala Asp Tyr Tyr Ile Ile Ala Gly Val Ile Tyr
 100 105 110
 Gln Ala Pro Asp Leu Gly Ser Val Ile Asn Ser Arg Val Leu Thr Ala
 115 120 125
 Val His Gly Ile Gln Ser Ala Phe Asp Glu Ala Met Ser Tyr Cys Arg
 130 135 140
 Tyr His Pro Ser Lys Gly Tyr Trp Trp His Phe Lys Asp His Glu Glu
 145 150 155 160
 Gln Asp Lys Val Arg Pro Lys Ala Lys Arg Lys Glu Glu Pro Ser Ser

0000250 "02129950"

```
<210> 40
<211> 245
<212> PRT
<213> Homo sapien
```

<210>	41
<211>	163
<212>	PRT

<400> 41

<210> 42

<211> 243

<212> PRT

<213> Homo sapien

<400> 42

Val 1	Asp	Ile	Arg	Asp 5	Asn	Leu	Leu	Gly	Ile 10	Ser	Trp	Val	Asp	Ser 15	Ser
Trp	Ile	Pro	Ile	Leu	Asn	Ser	Gly	Ser	Val	Leu	Asp	Tyr	Phe	Ser	Glu
			20					25					30		
Arg	Ser	Asn	Pro	Phe	Tyr	Asp	Arg	Thr	Cys	Asn	Asn	Glu	Val	Val	Lys
		35					40					45			
Met	Gln	Arg	Leu	Thr	Leu	Glu	His	Leu	Asn	Gln	Met	Val	Gly	Ile	Glu
	50					55					60				
Tyr	Ile	Leu	Leu	His	Ala	Gln	Glu	Pro	Ile	Leu	Phe	Ile	Ile	Arg	Lys
65					70					75				80	
Gln	Gln	Arg	Gln	Ser	Pro	Ala	Gln	Val	Ile	Pro	Leu	Ala	Asp	Tyr	Tyr
				85					90					95	
Ile	Ile	Ala	Gly	Val	Ile	Tyr	Gln	Ala	Pro	Asp	Leu	Gly	Ser	Val	Ile
			100					105					110		
Asn	Ser	Arg	Val	Leu	Thr	Ala	Val	His	Gly	Ile	Gln	Ser	Ala	Phe	Asp
		115					120					125			
Glu	Ala	Met	Ser	Tyr	Cys	Arg	Tyr	His	Pro	Ser	Lys	Gly	Tyr	Trp	Trp
	130					135					140				
His	Phe	Lys	Asp	His	Glu	Gln	Asp	Lys	Val	Arg	Pro	Lys	Ala	Lys	
145					150				155					160	
Arg	Lys	Glu	Glu	Pro	Ser	Ser	Ile	Phe	Gln	Arg	Gln	Arg	Val	Asp	Ala
				165					170					175	


```
<210> 45
<211> 324
<212> PRT
<213> Homo sapien
```

Arg 1	Arg	Pro	Val	Met 5	Ala	Gln	Glu	Thr	Ala 10	Pro	Pro	Cys	Gly	Pro 15	Val
Ser	Arg	Gly	Asp 20	Ser	Pro	Ile	Ile	Glu 25	Lys	Met	Glu	Lys	Arg 30	Thr	Cys
Ala	Leu	Cys 35	Pro	Glu	Gly	His	Glu 40	Trp	Ser	Gln	Ile	Tyr 45	Phe	Ser	Pro
Ser	Gly 50	Asn	Ile	Val	Ala	His 55	Glu	Asn	Cys	Leu	Leu 60	Tyr	Ser	Ser	Gly
Leu 65	Val	Glu	Cys	Glu	Thr 70	Leu	Asp	Leu	Arg 75	Asn	Thr	Ile	Arg	Asn	Phe 80
Asp	Val	Lys	Ser 85	Val	Lys	Lys	Glu	Ile 90	Trp	Arg	Gly	Arg	Arg	Leu 95	Lys
Cys	Ser	Phe 100	Cys	Asn	Lys	Gly	Gly 105	Ala	Thr	Val	Gly	Cys	Asp 110	Leu	Trp
Phe	Cys 115	Lys	Lys	Ser	Tyr	His 120	Tyr	Val	Cys	Ala	Lys	Lys 125	Asp	Gln	Ala
Ile	Leu 130	Gln	Val	Asp	Gly	Asn 135	His	Gly	Thr	Tyr	Lys 140	Leu	Phe	Cys	Pro
Glu 145	His	Ser	Pro	Glu	Gln 150	Glu	Glu	Ala	Thr	Glu	Ser 155	Ala	Asp	Asp	Pro 160
Ser	Met	Lys	Lys 165	Lys	Arg	Gly	Lys	Asn 170	Lys	Arg	Leu	Ser	Ser	Gly 175	Pro
Pro	Ala	Gln 180	Pro	Lys	Thr	Met	Lys	Cys 185	Ser	Asn	Ala	Lys	Arg 190	His	Met
Thr	Glu 195	Glu	Pro	His	Gly	His 200	Thr	Asp	Ala	Ala	Val 205	Lys	Ser	Pro	Phe
Leu	Lys 210	Lys	Cys	Gln	Glu	Ala 215	Gly	Leu	Leu	Thr	Glu 220	Leu	Phe	Glu	His
Ile 225	Leu	Glu	Asn	Met	Asp 230	Ser	Val	His	Gly	Arg 235	Leu	Val	Asp	Glu	Thr 240
Ala	Ser	Glu	Ser	Asp	Tyr	Glu	Gly	Ile	Glu	Thr	Leu	Leu	Phe	Asp	Cys

```
<210> 46
<211> 244
<212> PRT
<213> Homo sapien
```

<210>	47
<211>	14
<212>	DNA

gagtctgagc	agaaagtaaa	agcagccttg	gcagccacgt	tagaggaata	caaagccaca	60
gtggccagtg	accagataga	gatgaatcgc	ctgaaggctc	agctggagaa	tgaaaagcag	120
aaagtggcag	agctgtattc	tatccataac	tctggagaca	aatctgatat	tcaggacctc	180


```
<210> 52
<211> 491
<212> DNA
<213> Homo sapien
```

```
<210> 53
<211> 787
<212> DNA
<213> Homo sapien
```

```
<210> 54
<211> 386
<212> DNA
<213> Homo sapien
```


Ala Gln Leu Glu Asn Glu Lys Gln Lys Val Ala Glu Leu Tyr Ser Ile
 35 40 45
 His Asn Ser Gly Asp Lys Ser Asp Ile Gln Asp Leu Leu Glu Ser Val
 50 55 60
 Arg Leu Asp Lys Glu Lys Ala Glu Thr Leu Ala Ser Ser Leu Gln Glu
 65 70 75 80
 Asp Leu Ala His Thr Arg Asn Asp Ala Asn Arg Leu Gln Asp Ala Ile
 85 90 95
 Ala Lys Val Glu Asp Glu Tyr Arg Ala Phe Gln Glu Glu Ala Lys Lys
 100 105 110
 Gln Ile Glu Asp Leu Asn Met Thr Leu Glu Lys Leu Arg Ser Asp Leu
 115 120 125
 Asp Glu Lys Glu Thr Glu Arg Ser Asp Met Lys Glu Thr Ile Phe Glu
 130 135 140
 Leu Glu Asp Glu Val Glu Gln His Arg Ala Val Lys Leu His Asp Asn
 145 150 155 160
 Leu Ile Ile Ser Asp Leu Glu Asn Thr Val Lys Lys Leu Gln Asp Gln
 165 170 175
 Lys His Asp Met Glu Arg Glu Ile Lys Thr Leu His Arg Arg Leu Arg
 180 185 190
 Glu Glu Ser Ala Glu Trp Arg Gln Phe Gln Ala Asp Leu Gln Thr Ala
 195 200 205
 Val Val Ile Ala Asn Asp Ile Lys Ser Glu Ala Gln Glu Glu Ile Gly
 210 215 220
 Asp Leu Lys Arg Arg Leu His Glu Ala Gln Glu Lys Asn Glu Lys Leu
 225 230 235 240
 Thr Lys Glu Leu Glu Glu Ile Lys Ser Arg Lys Gln Glu Glu Glu Arg
 245 250 255
 Gly Gly Tyr

<210> 59
 <211> 125
 <212> PRT
 <213> Homo sapien

<400> 59

Gly Thr Ser Phe Ser Lys Asn His Ala Ala Pro Phe Ser Lys Val Leu
 1 5 10 15
 Thr Phe Tyr Arg Lys Glu Pro Phe Thr Leu Glu Ala Tyr Tyr Ser Ser
 20 25 30
 Pro Gln Asp Leu Pro Tyr Pro Asp Pro Ala Ile Ala Gln Phe Ser Val
 35 40 45
 Gln Lys Val Thr Pro Gln Ser Asp Gly Ser Ser Ser Lys Val Lys Val
 50 55 60
 Lys Val Arg Val Asn Val His Gly Ile Phe Ser Val Ser Ser Ala Ser
 65 70 75 80
 Leu Val Glu Val His Lys Ser Glu Glu Asn Glu Glu Pro Met Glu Thr
 85 90 95
 Asp Gln Asn Ala Lys Glu Glu Glu Lys Met Gln Val Asp Gln Glu Glu
 100 105 110
 Pro His Val Glu Glu Gln Gln Gln Thr Pro Gly Arg
 115 120 125

0000260"024350


```
<210> 62
<211> 418
<212> PRT
<213> Homo sapien
```

	<400>															62
Met 1	Tyr	Arg	Pro	Ala 5	Arg	Val	Thr	Ser	Thr 10	Ser	Arg	Phe	Leu 15	Asn 15	Pro	
Tyr	Val	Val	Cys 20	Phe	Ile	Val	Val	Ala 25	Gly	Val	Val	Ile	Leu 30	Ala	Val	
Thr	Ile	Ala 35	Leu	Leu	Val	Tyr	Phe 40	Leu	Ala	Phe	Asp 45	Gln	Lys	Ser	Tyr	
Phe	Tyr 50	Arg	Ser	Ser	Phe	Gln 55	Leu	Leu	Asn	Val	Glu 60	Tyr	Asn	Ser	Gln	
Leu 65	Asn	Ser	Pro	Ala	Thr 70	Gln	Glu	Tyr	Arg	Thr 75	Leu	Ser	Gly	Arg	Ile 80	
Glu	Ser	Leu	Ile	Thr 85	Lys	Thr	Phe	Lys	Glu 90	Ser	Asn	Leu	Arg	Asn	Gln	
Phe	Ile	Arg	Ala 100	His	Val	Ala	Lys	Leu 105	Arg	Gln	Asp	Gly	Ser	Gly	Val	
Arg	Ala	Asp 115	Val	Val	Met	Lys	Phe 120	Gln	Phe	Thr	Arg	Asn 125	Asn	Asn	Gly	
Ala	Ser 130	Met	Lys	Ser	Arg	Ile 135	Glu	Ser	Val	Leu	Arg 140	Gln	Met	Leu	Asn	
Asn 145	Ser	Gly	Asn	Leu	Glu 150	Ile	Asn	Pro	Ser	Thr 155	Glu	Ile	Thr	Ser	Leu 160	
Thr	Asp	Gln	Ala 165	Ala	Ala	Asn	Trp	Leu	Ile 170	Asn	Glu	Cys	Gly	Ala 175	Gly	
Pro	Asp	Leu 180	Ile	Thr	Leu	Ser	Glu 185	Gln	Arg	Ile	Leu	Gly 190	Gly	Thr	Glu	
Ala	Glu 195	Glu	Gly	Ser	Trp	Pro	Trp 200	Gln	Val	Ser	Leu	Arg 205	Leu	Asn	Asn	
Ala 210	His	His	Cys	Gly	Gly 215	Ser	Leu	Ile	Asn	Asn	Met 220	Trp	Ile	Leu	Thr	
Ala 225	Ala	His	Cys	Phe	Arg 230	Ser	Asn	Ser	Asn	Pro	Arg 235	Asp	Trp	Ile	Ala 240	
Thr	Ser	Gly	Ile 245	Ser	Thr	Thr	Phe	Pro	Lys 250	Leu	Arg	Met	Arg	Val	Arg	
Asn	Ile	Leu	Ile 260	His	Asn	Asn	Tyr	Lys 265	Ser	Ala	Thr	His	Glu	Asn	Asp	
Ile	Ala 275	Leu	Val	Arg	Leu	Glu	Asn 280	Ser	Val	Thr	Phe	Thr 285	Lys	Asp	Ile	
His	Ser 290	Val	Cys	Leu	Pro	Ala 295	Ala	Thr	Gln	Asn	Ile 300	Pro	Pro	Gly	Ser	
Thr 305	Ala	Tyr	Val	Thr	Gly 310	Trp	Gly	Ala	Gln	Glu 315	Tyr	Ala	Gly	His	Thr 320	

```
<210> 63
<211> 776
<212> DNA
<213> Homo sapien
```

```
<210> 64
<211> 160
<212> DNA
<213> Homo sapien
```

```
<210> 65
<211> 72
<212> PRT
<213> Homo sapien
```

<400> 65
Leu Ser Ala Met Gly Phe Thr Ala Ala Gly Ile Ala Ser Ser Ser Ile
1 5 10 15
Ala Ala Lys Met Met Ser Ala Ala Ala Ile Ala Asn Gly Gly Gly Val

	20		25		30										
Ala	Ser	Gly	Ser	Leu	Val	Ala	Thr	Leu	Gln	Ser	Leu	Gly	Ala	Thr	Gly
	35		40		45										
Leu	Ser	Gly	Leu	Thr	Lys	Phe	Ile	Leu	Gly	Ser	Ile	Gly	Ser	Ala	Ile
	50		55		60										
Ala	Ala	Val	Ile	Ala	Arg	Phe	Tyr								
65			70												

<210> 66
 <211> 2581
 <212> DNA
 <213> Homo sapien

<400> 66

ctttcaaccc	gcgctcgccg	gctccagccc	cgcgcgcccc	cacccttgc	cctcccggcg	60
gctccgcagg	gtgaggtggc	tttgaccccc	ggttgccccg	ccagcacgac	cgaggaggtg	120
gctggacagc	tggaggatga	acggagaagc	cgactgcccc	acagacctgg	aaatggccgc	180
cccaaaggc	caagaccgtt	ggtcccagga	agacatgctg	actttgctgg	aatgcatgaa	240
gaacaacctt	ccatccaatg	acagctccaa	gttcaaaacc	accgaatcac	acatggactg	300
ggaaaaagta	gcatttaaag	acttttcttg	agacatgtgc	aagctcaaat	gggtggagat	360
ttctaattgag	gtgaggaagt	tccgtacatt	gacagaattg	atcctcgatg	ctcaggaaca	420
tgttaaaaaat	ccttacaaag	gcaaaaaact	caagaaacac	ccagacttcc	caaagaagcc	480
cctgaccctt	tatttccgct	tcttcatgga	gaagcggggc	aagtatgcga	aactccacc	540
tgagatgagc	aacctggacc	taaccaagat	tctgtccaag	aaatacaagg	agcttccgga	600
gaagaagaag	atgaaatata	ttcaggactt	ccagagagag	aaacaggagt	tcgagcgaaa	660
cctggccccg	ttcaggggag	atcacccccg	cctaattccag	aatgccaaag	aatcgacat	720
cccagagaag	cccaaaaccc	cccagcagct	gtggtacacc	cacgagaaga	aggtgtatct	780
caaagtgcg	ccagatgcc	ctacgaagg	ggtgaaggac	tccctgggga	agcagtggtc	840
tcagctctcg	gacaaaaaga	ggctgaaatg	gattcataag	gccctggagc	agcgggaagg	900
gtacgaggag	atcatgagag	actatatcca	gaagcaccca	gagctgaaca	tcagtggagg	960
gggtatcacc	aagtccaccc	tcaccaaggc	cgaacgccag	ctcaaggaca	agtttgacgg	1020
gcgacccacc	aagccacctc	cgaacagcta	ctcgtgttac	tgcgcagagc	tcattggcaa	1080
catgaaggac	gtgccagca	cagagcgcat	ggtgctgtgc	agccagcagt	ggaagctgct	1140
gtcccagaag	gagaaggacg	cctatcacaa	gaagtgtgat	cagaaaaaga	aagattacga	1200
ggtggagctg	ctcgtttcc	tcgagagcct	gcctgaggag	gagcagcagc	gggtcttggg	1260
ggaagagaag	atgctgaaca	tcaacaagaa	gcaggccacc	agccccgcct	ccaagaagcc	1320
agcccaggaa	gggggcaagg	gcggctccga	gaagcccaag	cggccctgtg	cggccatgtt	1380
catcttctcg	gaggagaaac	ggcggcagct	gcaggaggag	cggcctgagc	tctccgagag	1440
cgagctgacc	cgctgctgg	cccgaatgtg	gaacgacctg	tctgagaaga	agaaggccaa	1500
gtacaaggcc	cgagaggcgg	cgctcaaggc	tcagtggag	aggaagcccg	gcggggagcg	1560
cgaggaacgg	ggcaagctgc	ccgagtcctc	caaaagagct	gaggagatct	ggcaacagag	1620
cgttatcggc	gactacctgg	cccgttcaa	gaatgaccgg	gtgaaggcct	tgaaagccat	1680
ggaaatgacc	tggaataaca	tggaaaagaa	ggagaaactg	atgtggatta	agaaggcagc	1740
cgaagaccaa	aagcgatatg	agagagagct	gagtgagatg	cgggcacctc	cagctgctac	1800
aaattcttcc	aagaagatga	aattccaggg	agaacccaag	aagcctccca	tgaacggtta	1860
ccagaagttc	tcccaggagc	tgtgtccaa	tggggagctg	aaccacctgc	cgctgaagga	1920
gcgcatggtg	gagatcgcca	gtcgtggcca	gcgcatctcc	cagagccaga	aggagcacta	1980
caaaaagctg	gccgaggagc	agcaaaagca	gtacaagggtg	cacctggacc	tctgggttaa	2040
gagcctgtct	ccccaggacc	gtgcagcata	taaagagtac	atctccaata	aacgtaagag	2100
catgaccaag	ctgcgaggcc	caaaccccaa	atccagccgg	actactctgc	agtccaagtc	2160
ggagtccgag	gaggatgatg	aagaggatga	ggatgacgag	gacgaggatg	aagaagagga	2220
agatgatgag	aatggggact	cctctgaaga	tggcggcgac	tcctctgagt	ccagcagcga	2280
ggacgagagc	gaggatgggg	atgagaatga	agaggatgac	gaggacgaag	acgacgacga	2340
ggatgacgat	gaggatgaag	ataatgagtc	cgagggcagc	agctccagct	cctctctctt	2400

0065710 "092000

			340				345				350				
Asp	Tyr	Glu	Val	Glu	Leu	Leu	Arg	Phe	Leu	Glu	Ser	Leu	Pro	Glu	Glu
			355				360				365				
Glu	Gln	Gln	Arg	Val	Leu	Gly	Glu	Glu	Lys	Met	Leu	Asn	Ile	Asn	Lys
			370				375				380				
Lys	Gln	Ala	Thr	Ser	Pro	Ala	Ser	Lys	Lys	Pro	Ala	Gln	Glu	Gly	Gly
385				390				395				400			
Lys	Gly	Gly	Ser	Glu	Lys	Pro	Lys	Arg	Pro	Val	Ser	Ala	Met	Phe	Ile
			405				410				415				
Phe	Ser	Glu	Glu	Lys	Arg	Arg	Gln	Leu	Gln	Glu	Glu	Arg	Pro	Glu	Leu
			420				425				430				
Ser	Glu	Ser	Glu	Leu	Thr	Arg	Leu	Leu	Ala	Arg	Met	Trp	Asn	Asp	Leu
			435				440				445				
Ser	Glu	Lys	Lys	Lys	Ala	Lys	Tyr	Lys	Ala	Arg	Glu	Ala	Ala	Leu	Lys
			450				455				460				
Ala	Gln	Ser	Glu	Arg	Lys	Pro	Gly	Gly	Glu	Arg	Glu	Glu	Arg	Gly	Lys
465				470				475				480			
Leu	Pro	Glu	Ser	Pro	Lys	Arg	Ala	Glu	Glu	Ile	Trp	Gln	Gln	Ser	Val
			485				490				495				
Ile	Gly	Asp	Tyr	Leu	Ala	Arg	Phe	Lys	Asn	Asp	Arg	Val	Lys	Ala	Leu
			500				505				510				
Lys	Ala	Met	Glu	Met	Thr	Trp	Asn	Asn	Met	Glu	Lys	Lys	Glu	Lys	Leu
			515				520				525				
Met	Trp	Ile	Lys	Lys	Ala	Ala	Glu	Asp	Gln	Lys	Arg	Tyr	Glu	Arg	Glu
			530				535				540				
Leu	Ser	Glu	Met	Arg	Ala	Pro	Pro	Ala	Ala	Thr	Asn	Ser	Ser	Lys	Lys
545				550				555				560			
Met	Lys	Phe	Gln	Gly	Glu	Pro	Lys	Lys	Pro	Pro	Met	Asn	Gly	Tyr	Gln
			565				570				575				
Lys	Phe	Ser	Gln	Glu	Leu	Leu	Ser	Asn	Gly	Glu	Leu	Asn	His	Leu	Pro
			580				585				590				
Leu	Lys	Glu	Arg	Met	Val	Glu	Ile	Gly	Ser	Arg	Trp	Gln	Arg	Ile	Ser
			595				600				605				
Gln	Ser	Gln	Lys	Glu	His	Tyr	Lys	Lys	Leu	Ala	Glu	Glu	Gln	Gln	Lys
			610				615				620				
Gln	Tyr	Lys	Val	His	Leu	Asp	Leu	Trp	Val	Lys	Ser	Leu	Ser	Pro	Gln
625				630				635				640			
Asp	Arg	Ala	Ala	Tyr	Lys	Glu	Tyr	Ile	Ser	Asn	Lys	Arg	Lys	Ser	Met
			645				650				655				
Thr	Lys	Leu	Arg	Gly	Pro	Asn	Pro	Lys	Ser	Ser	Arg	Thr	Thr	Leu	Gln
			660				665				670				
Ser	Lys	Ser	Glu	Ser	Glu	Glu	Asp	Asp	Glu	Glu	Asp	Glu	Asp	Asp	Glu
			675				680				685				
Asp	Glu	Asp	Glu	Glu	Glu	Glu	Asp	Asp	Glu	Asn	Gly	Asp	Ser	Ser	Glu
			690				695				700				
Asp	Gly	Gly	Asp	Ser	Ser	Glu	Ser	Ser	Ser	Glu	Asp	Glu	Ser	Glu	Asp
705				710				715				720			
Gly	Asp	Glu	Asn	Glu	Glu	Asp	Asp	Glu	Asp	Glu	Asp	Asp	Asp	Glu	Asp
			725				730				735				
Asp	Asp	Glu	Asp	Glu	Asp	Asn	Glu	Ser	Glu	Gly	Ser	Ser	Ser	Ser	Ser
			740				745				750				
Ser	Ser	Leu	Gly	Asp	Ser	Ser	Asp	Phe	Asp	Ser	Asn				
			755				760								

<210> 68
 <211> 434
 <212> DNA
 <213> Homo sapien

<400> 68
 ctaagatgct ggatgctgaa gacatcgctc gaactgcccg gccagatgag aaagccatta 60
 tgacttatgt gtctagcttc tatcatgcct tctctggagc ccagaaggca gaaacagcag 120
 ccaatcgcat ctgcaaagtg ttggcgggtca atcaagagaa cgagcagctt atggaagact 180
 atgagaagct ggccagtgat ctgttggagt ggatccgccg caccatccca tggctggaga 240
 atcgggtgcc tgagaacacc atgcatgcc a tgcagcagaa gctggaggac ttccgagact 300
 atagacgcct gcacaagccg cccaaggtgc aggagaagtg ccagctggag atcaacttta 360
 acacgctgca gaccaaactg cggctcagca accggcctgc cttcatgccc tccgagggca 420
 ggatgggtctc ggat 434

<210> 69
 <211> 244
 <212> DNA
 <213> Homo sapien

<400> 69
 aggcagcatg ctcggtgaga gtcatcacca ctccctaata tcaagtacgc agggacacaa 60
 aactgcgga aggccgcagg gtccctctgcc taggaaaacc agagaccttt gttcacttgt 120
 ttatgtgctg accttccctc cactattgtc ctgtgacct gccaaatccc cttttgtgag 180
 aaacacccaa gaatgatcaa taaaaaataa attaatttag gaaaaaaaaa aaaaaaaact 240
 cgag 244

<210> 70
 <211> 437
 <212> DNA
 <213> Homo sapien

<400> 70
 ctgggacggg agcgtccagc gggactcgaa ccccagatgt gaaggcggtt ctggaaagtc 60
 cttggtccct ggatccagcg tgggccagcc cagagcccggt gccgcacatc cttgcgtcct 120
 ccaggcagtg ggaccccgcg agctgcacgt ccctgggcac ggacaagtgt gaggcactgt 180
 tggggctgtg ccagggtgcg ggtgggctgc cccctttctc agaaccttcc agcctggtgc 240
 cgtggccccc aggccggagt ctctctaagg ctgtgaggcc acccctgtcc tggcctccgt 300
 tctcgagca gcagaccttg cccgtgatga gcggggaggc ccttggtctg ctggggccagg 360
 ctggttccct ggccatggg gctgcacctc tgggggagcc agccaaggag gaccccatgc 420
 tggcgagga agccggg 437

<210> 71
 <211> 271
 <212> DNA
 <213> Homo sapien

<400> 71
 gcgcagagtt ctgtcgtcca ccatcgagtg aggaagagag cattggttcc cctgagatag 60
 aagagatggc tctcttcagt gccagctctc catacattaa cccgatcatc ccctttactg 120
 gaccaatcca aggagggctg caggagggac ttcagggtgac cctccagggg actaccgaga 180
 gttttgcaca aaagtgtgtg gtgaactttt cagaacagct tcaatggaga tgacttggcc 240
 ttccacttca accccgggta tgaggaagga g 271

<213> Homo sapien

[illegible]

<211> 69

<213> Homo sapien

[illegible]

<211> 96

<213> Homo sapien

Glu	Pro	Tyr	Pro	Glu	Val	Ser	Arg	Ile	Pro	Thr	Val	Arg	Gly	Cys	Asn
1				5					10					15	
Gly	Ser	Leu	Ser	Gly	Ala	Leu	Ser	Cys	Cys	Glu	Asp	Ser	Ala	Gln	Gly
			20					25					30		
Ser	Gly	Pro	Pro	Lys	Ala	Pro	Thr	Val	Ala	Glu	Gly	Pro	Ser	Ser	Cys
		35					40					45			

Leu Arg Arg Asn Val Ile Ser Glu Arg Glu Arg Arg Lys Arg Met Ser
 50 55 60
 Leu Ser Cys Glu Arg Leu Arg Ala Leu Leu Pro Gln Phe Asp Gly Arg
 65 70 75 80
 Arg Glu Asp Met Ala Ser Val Leu Glu Met Ser Val Ala Ile Pro Ala
 85 90 95

<210> 78
 <211> 2076
 <212> DNA
 <213> Homo sapien

<400> 78

agaaaaagtc tatgttttgc gaaatacaga tccaagacaa agacaggatg ggcactgctg 60
 gaaaagtatt taaatgcaaa gcagctgtgc tttgggagca gaagcaaccc ttctccattg 120
 aggaaataga agttgcccc ccaaagacta aagaagttcg cattaagatt ttggccacag 180
 gaatctgtcg cacagatgac catgtgataa aaggaacaat ggtgtccaag tttccagtga 240
 ttgtgggaca tgaggcaact gggattgtag agagcattgg agaaggagtg actacagtga 300
 aaccagggtga caaagtcata cctctctttc tgccacaatg tagagaatgc aatgcttgct 360
 gcaacccaga tggcaacctt tgcattagga gcgatattac tggctgtgga gtactggctg 420
 atggcaccac cagattttaca tgcaagggca aaccagtcca ccacttcatg aacaccagta 480
 catttaccga gtacacagtg gtggatgaat cttctgttgc taagattgat gatgcagctc 540
 ctcttgagaa agtctgttta attggctgtg gggtttccac tggatatggc gctgctgtta 600
 aaactggcaa ggtcaaacct gggtccactt gcgtcgtctt tggcctgaga ggagttggcc 660
 tgtcagtcac catgggctgt aagtcagctg gtgcatctag gatcattggg attgacctca 720
 acaaaagcaa atttgagaag gccatggctg taggtgccac tgagtgtatc agtcccaagg 780
 actctaccaa acccatcagt gaggtgctgt cagaaatgac aggcaacaac gtgggataca 840
 cctttgaagt tattgggcat cttgaaacca tgattgatgc cctggcatcc tgccacatga 900
 actatgggac cagcgtgggt gtaggagttc ctccatcagc caagatgctc acctatgacc 960
 cgatgttgct cttcactgga cgcacatgga agggatgtgt ctttggaggt ttgaaaagca 1020
 gagatgatgt cccaaaacta gtgactgagt tcctggcaaa gaaatttgac ctggaccagt 1080
 tgataactca tgtcttacca tttaaaaaaa tcagtgaagg atttgagctg ctcaattcag 1140
 gacaaagcat tcgaacggtc ctgacgtttt gagatccaaa gtggcaggag gtctgtgttg 1200
 tcatggtgaa ctggagtttc tcttgtgaga gttccctcat ctgaaatcat gtatctgtct 1260
 cacaaatata agcataagta gaagattttgt tgaagacata gaacccttat aaagaattat 1320
 taacctttat aaacatttaa agtcttgtga gcacctggga attagtataa taacaatgtt 1380
 aatatttttg atttacattt tgtaaggcta taattgtatc ttttaagaaa acatacactt 1440
 ggatttctat gttgaaatgg agatttttaa gagttttaac cagctgctgc agatatatat 1500
 ctcaaaacag atatagcgtt taaagatata gtaaattgcat ctctagagt aatattcact 1560
 taacacattg aaactattat tttttagatt tgaatataaa tgtatttttt aaacacttgt 1620
 tatgagttaa cttggattac attttgaaat cagttcattc catgatgcat attactggat 1680
 tagattaaga aagacagaaa agattaaggg acgggcacat ttttcaacga ttaagaatca 1740
 tcattacata acttggtgaa actgaaaaag tatatcatat gggtagacaa ggctatttgc 1800
 cagcatatat taatatatta gaaaatattc cttttgtaat actgaatata aacatagagc 1860
 tagaatcata ttatcatact tatcataatg ttcaatttga tacagtagaa ttgcaagtcc 1920
 ttaagtcctt attcactgtg cttagtagtg actccattta ataaaaagtg tttttagttt 1980
 ttaacaacta cactgatgta tttatatata tttataacat gttaaaaatt ttttaaggaaa 2040
 ttaaaaaatta tataaaaaaa aaaaaaaaaa ctcgag 2076

<210> 79
 <211> 2790
 <212> DNA
 <213> Homo sapien

85 90 95
 Phe Ile Arg Ala His Val Ala Lys Leu Arg Gln Asp Gly Ser Gly Val
 100 105 110
 Arg Ala Asp Val Val Met Lys Phe Thr Arg Asn Asn Gly
 115 120 125
 Ala Ser Met Lys Ser Arg Ile Glu Ser Val Leu Arg Gln Met Leu Asn
 130 135 140
 Asn Ser Gly Asn Leu Glu Ile Asn Pro Ser Thr Glu Ile Thr Ser Leu
 145 150 155 160
 Thr Asp Gln Ala Ala Ala Asn Trp Leu Ile Asn Glu Cys Gly Ala Gly
 165 170 175
 Pro Asp Leu Ile Thr Leu Ser Glu Gln Arg Ile Leu Gly Gly Thr Glu
 180 185 190
 Ala Glu Glu Gly Ser Trp Pro Trp Gln Val Ser Leu Arg Leu Asn Asn
 195 200 205
 Ala His His Cys Gly Gly Ser Leu Ile Asn Asn Met Trp Ile Leu Thr
 210 215 220
 Ala Ala His Cys Phe Arg Ser Asn Ser Asn Pro Arg Asp Trp Ile Ala
 225 230 235 240
 Thr Ser Gly Ile Ser Thr Thr Phe Pro Lys Leu Arg Met Arg Val Arg
 245 250 255
 Asn Ile Leu Ile His Asn Asn Tyr Lys Ser Ala Thr His Glu Asn Asp
 260 265 270
 Ile Ala Leu Val Arg Leu Glu Asn Ser Val Thr Phe Thr Lys Asp Ile
 275 280 285
 His Ser Val Cys Leu Pro Ala Ala Thr Gln Asn Ile Pro Pro Gly Ser
 290 295 300
 Thr Ala Tyr Val Thr Gly Trp Gly Ala Gln Glu Tyr Ala Gly His Thr
 305 310 315 320
 Val Pro Glu Leu Arg Gln Gly Gln Val Arg Ile Ile Ser Asn Asp Val
 325 330 335
 Cys Asn Ala Pro His Ser Tyr Asn Gly Ala Ile Leu Ser Gly Met Leu
 340 345 350
 Cys Ala Gly Val Pro Gln Gly Gly Val Asp Ala Cys Gln Gly Asp Ser
 355 360 365
 Gly Gly Pro Leu Val Gln Glu Asp Ser Arg Arg Leu Trp Phe Ile Val
 370 375 380
 Gly Ile Val Ser Trp Gly Asp Gln Cys Gly Leu Pro Asp Lys Pro Gly
 385 390 395 400
 Val Tyr Thr Arg Val Thr Ala Tyr Leu Asp Trp Ile Arg Gln Gln Thr
 405 410 415
 Gly Ile

<210> 84
 <211> 489
 <212> DNA
 <213> Homo sapien

<400> 84

aaaagggttaa gcttgatgat taccaggaac gaatgaacaa aggggaaagg cttaatcaag	60
atcagctgga tgccgtttct aagtaaccagg aagtcacaaa taatttgag ttgcaaaa	120
aattacagag gagtttcatg gcactaagtc aagatattca gaaaacaata aagaagacag	180
cacgtcggga gcagcttatg agagaagaag ctgaacagaa acgtttaaaa actgtacttg	240

agctacagta	tgttttggac	aaattgggag	atgatgaagt	gcggactgac	ctgaaacaag	300
gtttgaatgg	agtgccaaata	ttgtccgaag	aggagttgtc	attggttgat	gaattctata	360
agctagtaga	ccctgaacgg	gacatgagct	tgaggttgaa	tgaacagtat	gaacatgcct	420
ccattcacct	gtgggacctg	ctggaaggga	aggaaaaacc	tgtatgtgga	accacctata	480
aagttctaa						489

<210> 85

<211> 304

<212> DNA

<213> Homo sapien

<400> 85

gggacctgga	ggaggccacg	ctgcagcatg	aagccacagc	agccaccctg	aggaagaagc	60
acgcggacag	cgtggccgag	ctcggggagc	agatcgacaa	cctgcagcgg	gtgaagcaga	120
agctggagaa	ggagaagagc	gagatgaaga	tggagatcga	tgacctcgct	tgtaacatgg	180
aggtcacatc	caaataaag	ggaaaccttg	agaagatgtg	ccgcacactg	gaggaccaag	240
tgagtgaagt	gaagaccag	gaggaggaac	agcagcggct	gatcaatgaa	ctgactgcgc	300
agag						304

<210> 86

<211> 296

<212> DNA

<213> Homo sapien

<400> 86

gaaaatcctt	cctttgaatg	ggaatctcca	agcagttgaa	ttgggcgaaa	aaagaacctc	60
ttccttaagg	attaaaaatg	ttagggcaac	acgtgttact	tccacttcca	gattttctgaa	120
tccatatgtt	gtatgtttcc	ttgtcctccc	aggggttgtg	atcctggcag	tcccatagc	180
tctacttggt	taactttttag	cttttgatca	aaaatcttac	ttttattgga	gcaattttcc	240
actcccaa	ggtgaatata	atagtccgtt	taattccccc	gcttcaccgg	gaattc	296

<210> 87

<211> 904

<212> DNA

<213> Homo sapien

<400> 87

gtgtccagga	aacgattcat	gaacataaca	agcttgctgc	aaattcagat	catctcatgc	60
agattcaaaa	atgtgagttg	gtcttgatcc	acacctaccc	agttggtgaa	gacagccttg	120
tatctgatcg	ttctaaaaaa	gagttgtccc	cggttttaac	cagtgaagtt	catagtgttc	180
gtgcaggacg	gcactcttgc	accaaattga	atattttagt	acagcaacat	tttgacttgg	240
cttcaactac	tattacaaat	attccaatga	aggaagaaca	gcatgctaac	acatctgcca	300
attatgatgt	ggagctactt	catcacaaag	atgcacatgt	agattttcctg	aaaagtgggtg	360
attcgcacat	aggtggcggc	agtcgagaag	gctcgtttaa	agaaacaata	acattaaagt	420
ggtgtacacc	aaggacaaat	aacattgaat	tacactattg	tactggagct	tatcggattt	480
cacctgtaga	tgtaaatagt	agaccttcct	cctgccttac	taattttctt	ctaaatggtc	540
gttctgtttt	attggaacaa	ccacgaaagt	caggttctaa	agtcattagt	catatgctta	600
gtagccatgg	aggagagatt	tttttgacg	tccttagcag	ttctcgatcc	attctagaag	660
atccaccttc	aattagttaa	ggatgtggag	gaagagttac	agactaccgg	attacagatt	720
ttggtgaatt	tatgagggga	aaacagatta	actccttttc	tacaccccag	atataaaatc	780
gatggaagtc	ttgaggtccc	tttggaaccg	agccaaaaga	tcagttaaaa	aaacataccc	840
gttactggcc	tatgatttca	aaaaccacc	atttttaaca	tgcaagcggg	agttccgtta	900
acca						904

<210> 88
 <211> 387
 <212> DNA
 <213> Homo sapien

<400> 88
 cgtctctccc ccagtttgcc gttcacccgg agcgctcggg acttgccgat agtgggtgacg 60
 gcggcaacat gtctgtggct ttgcgcggccc cgaggcagcg aggcaagggg gagatcactc 120
 ccgctgcgat tcagaagatg ttggatgaca ataaccatct tattcagtgt ataattggact 180
 ctcagaataa aggaaaagacc tcagagtgtt ctcagtatca gcagatgttg cacacaaact 240
 tgggtatacct tgctacaata gcagattcta atcaaaaatat gcagtctctt ttaccagcac 300
 caccacacac gaatatgcct atgggtcctg gaggggatgaa tcagagcggg cctccccccac 360
 ctccacgctc tcacaacatg ccttcaa 387

<210> 89
 <211> 481
 <212> DNA
 <213> Homo sapien

<400> 89
 tgttcttggg cctgcggtgc tatagagcag gctcttctag gttggcagtt gccatggaat 60
 ctggacccaa aatgttggcc cccgtttgcc tgggtgaaaa taacaatgag cagctatttg 120
 tgaaccagca agctatacag attcttgaaa agattttctca gccagtgggtg gtgggtggcca 180
 ttgtaggact gtaccgtaca gggaaatcct acttgatgaa ccatctggca ggacagaatc 240
 atggcttccc tctgggctcc acggtgcagt ctgaaaccaa gggcatctgg atgtggtgcg 300
 tgccccaccc atccaagcca aaccacacccc tggctccttct ggacaccgaa ggtctgggcg 360
 atgtggaaaa gggtgaccct aagaatgact cctggatctt tgccctggct gtgctcctgt 420
 gcagcacctt tgtctacaac agcatgagca ccatcaacca ccaggccctg gagcagctgc 480
 a 481

<210> 90
 <211> 491
 <212> DNA
 <213> Homo sapien

<400> 90
 tgaaaactgt tcttggacct gcggtgctat agagcagggtt ggcagttgcc atggaatctg 60
 gacccaaaat gttggccccc gtttgccctg tggaaaataa caatgagcag ctattggtga 120
 accagcaagc tatacagatt cttgaaaaga tttctcagcc agtgggtgggtg gtggccattg 180
 taggactgta ccgtacaggg aaatcctact tgatgaacca tctggcagga cagaatcatg 240
 gcttccctct gggctccacg gtgcagctctg aaaccaagggt catctggatg tgggtgcgtgc 300
 cccaccatc caagccaaac cacaccctgg tccttctgga caccgaagggt ctgggcgatg 360
 tggaaaagggt tgaccctaag aatgactcct ggatctttgc cctggctgtg ctctgtgca 420
 gcacctttgt ctacaacagc atgagcacca tcaaccacca agccctggag cagctgcatt 480
 atgtgacgga c 491

<210> 91
 <211> 488
 <212> DNA
 <213> Homo sapien

<400> 91
 ttcgacagtc agccgcatct tcttttgcgt cggcagccga gccacatcgc tcagacacca 60
 tggggaagggt gaaggtcgga gtcaacggat ttggtcgtat tgggcgcctg gtcaccagggt 120

```

ctgcttttaa ctctggtaaa gtggatattg ttgccatcaa tgaccccttc attgacctca 180
actacatggt ttacatgttc caatatgatt ccacccatgg caaattccat ggcaccgtcg 240
aggetgagaa cgggaagctt gtcatcaatg gaaatcccat caccatcttc caggagcgag 300
atccctccaa aatcaagtgg ggcgatgctg gcgctgagta cgtcgtggag tccactggcg 360
tcttcaccac catggagaag gctggggctc atttgcaggg gggagccaaa agggtcatca 420
tctctgcccc tctgctgatg ccccatgttc gtcatgggtg tgaacccatga gaagtatgac 480
acagcctc

```

```

<210> 92
<211> 384
<212> DNA
<213> Homo sapien

```

```

<400> 92
gacagtcagc cgcattcttct tttgcgtcgc cagccgagcc acatcgctca gacaccatgg 60
ggaagggtgaa ggtcggagtc aacggatttg gtcgtattgg gcgcctggtc accagggctg 120
cttttaactc tggtaaagtg gatattgttg ccatcaatga ccccttcatt gacctcaact 180
acatgggttta catgttccaa tatgattcca cccatggcaa attccatggc accgtcgagg 240
ctgagaacgg gaagcttgct atcaatggaa atcccatcac catcttccag gacgcgagtc 300
cctccaaaat caagtggggc gatactggcg ctgagtacgt cgtggagtc actggcgtct 360
tcaccacccat ggagaaggct gggg
384

```

```

<210> 93
<211> 162
<212> PRT
<213> Homo sapien

```

```

<400> 93
Lys Gly Lys Leu Asp Asp Tyr Gln Glu Arg Met Asn Lys Gly Glu Arg
 1          5          10          15
Leu Asn Gln Asp Gln Leu Asp Ala Val Ser Lys Tyr Gln Glu Val Thr
          20          25          30
Asn Asn Leu Glu Phe Ala Lys Glu Leu Gln Arg Ser Phe Met Ala Leu
          35          40          45
Ser Gln Asp Ile Gln Lys Thr Ile Lys Lys Thr Ala Arg Arg Glu Gln
          50          55          60
Leu Met Arg Glu Glu Ala Glu Gln Lys Arg Leu Lys Thr Val Leu Glu
65          70          75          80
Leu Gln Tyr Val Leu Asp Lys Leu Gly Asp Asp Glu Val Arg Thr Asp
          85          90          95
Leu Lys Gln Gly Leu Asn Gly Val Pro Ile Leu Ser Glu Glu Glu Leu
          100          105          110
Ser Leu Leu Asp Glu Phe Tyr Lys Leu Val Asp Pro Glu Arg Asp Met
          115          120          125
Ser Leu Arg Leu Asn Glu Gln Tyr Glu His Ala Ser Ile His Leu Trp
          130          135          140
Asp Leu Leu Glu Gly Lys Glu Lys Pro Val Cys Gly Thr Thr Tyr Lys
145          150          155          160
Val Leu

```

```

<210> 94
<211> 100
<212> PRT

```

09557170.0955000

<213> Homo sapien

<400> 94

```

Asp Leu Glu Glu Ala Thr Leu Gln His Glu Ala Thr Ala Ala Thr Leu
 1          5          10          15
Arg Lys Lys His Ala Asp Ser Val Ala Glu Leu Gly Glu Gln Ile Asp
          20          25          30
Asn Leu Gln Arg Val Lys Gln Lys Leu Glu Lys Glu Lys Ser Glu Met
          35          40          45
Lys Met Glu Ile Asp Asp Leu Ala Cys Asn Met Glu Val Ile Ser Lys
          50          55          60
Ser Lys Gly Asn Leu Glu Lys Met Cys Arg Thr Leu Glu Asp Gln Val
          65          70          75          80
Ser Glu Leu Lys Thr Gln Glu Glu Glu Gln Gln Arg Leu Ile Asn Glu
          85          90          95
Leu Thr Ala Gln
          100

```

<210> 95

<211> 99

<212> PRT

<213> Homo sapien

<400> 95

```

Lys Ile Leu Pro Leu Asn Gly Asn Leu Gln Ala Val Glu Leu Gly Glu
 1          5          10          15
Lys Arg Thr Ser Ser Leu Arg Ile Lys Met Phe Arg Ala Thr Arg Val
          20          25          30
Thr Ser Thr Ser Arg Phe Leu Asn Pro Tyr Val Val Cys Phe Leu Val
          35          40          45
Leu Pro Gly Val Val Ile Leu Ala Val Pro Ile Ala Leu Leu Val Tyr
          50          55          60
Phe Leu Ala Phe Asp Gln Lys Ser Tyr Phe Tyr Trp Ser Asn Phe Pro
          65          70          75          80
Leu Pro Asn Val Glu Tyr Asn Ser Pro Phe Asn Ser Pro Ala Ser Pro
          85          90          95
Gly Ile Pro

```

<210> 96

<211> 257

<212> PRT

<213> Homo sapien

<400> 96

```

Val Gln Glu Thr Ile His Glu His Asn Lys Leu Ala Ala Asn Ser Asp
 1          5          10          15
His Leu Met Gln Ile Gln Lys Cys Glu Leu Val Leu Ile His Thr Tyr
          20          25          30
Pro Val Gly Glu Asp Ser Leu Val Ser Asp Arg Ser Lys Lys Glu Leu
          35          40          45
Ser Pro Val Leu Thr Ser Glu Val His Ser Val Arg Ala Gly Arg His
          50          55          60
Leu Ala Thr Lys Leu Asn Ile Leu Val Gln Gln His Phe Asp Leu Ala

```

000000"02729350

```
<210> 97
<211> 128
<212> PRT
<213> Homo sapien
```

```
<210> 98
<211> 159
<212> PRT
<213> Homo sapien
```


<400> 100
 Met Gly Lys Val Lys Val Gly Val Asn Gly Phe Gly Arg Ile Gly Arg
 1 5 10 15
 Leu Val Thr Arg Ala Ala Phe Asn Ser Gly Lys Val Asp Ile Val Ala
 20 25 30
 Ile Asn Asp Pro Phe Ile Asp Leu Asn Tyr Met Val Tyr Met Phe Gln
 35 40 45
 Tyr Asp Ser Thr His Gly Lys Phe His Gly Thr Val Glu Ala Glu Asn
 50 55 60
 Gly Lys Leu Val Ile Asn Gly Asn Pro Ile Thr Ile Phe Gln Glu Arg
 65 70 75 80
 Asp Pro Ser Lys Ile Lys Trp Gly Asp Ala Gly Ala Glu Tyr Val Val
 85 90 95
 Glu Ser Thr Gly Val Phe Thr Thr Met Glu Lys Ala Gly Ala His Leu
 100 105 110
 Gln Gly Gly Ala Lys Arg Val Ile Ile Ser Ala Pro
 115 120

<210> 101
 <211> 127
 <212> PRT
 <213> Homo sapien

<400> 101
 Gln Ser Ala Ala Ser Ser Phe Ala Ser Pro Ala Glu Pro His Arg Ser
 1 5 10 15
 Asp Thr Met Gly Lys Val Lys Val Gly Val Asn Gly Phe Gly Arg Ile
 20 25 30
 Gly Arg Leu Val Thr Arg Ala Ala Phe Asn Ser Gly Lys Val Asp Ile
 35 40 45
 Val Ala Ile Asn Asp Pro Phe Ile Asp Leu Asn Tyr Met Val Tyr Met
 50 55 60
 Phe Gln Tyr Asp Ser Thr His Gly Lys Phe His Gly Thr Val Glu Ala
 65 70 75 80
 Glu Asn Gly Lys Leu Val Ile Asn Gly Asn Pro Ile Thr Ile Phe Gln
 85 90 95
 Glu Arg Asp Pro Ser Lys Ile Lys Trp Gly Asp Thr Gly Ala Glu Tyr
 100 105 110
 Val Val Glu Ser Thr Gly Val Phe Thr Thr Met Glu Lys Ala Gly
 115 120 125

<210> 102
 <211> 1225
 <212> DNA
 <213> Homo sapien

<400> 102
 atggcgggcg ggtcgctgctc ggggggtggcg gcggcagagg gggcgggcg cctggcgggca 60
 gcggagacgg cagccgtgac ggtggcagcg gcggcgcggg acctgggcct gggggaatga 120
 ggcggcccg gcgggccagc ggcggagccg tgtagcggag aagctcccc tccctgcttc 180
 ccttgccga gccggggcg cgcgcgcacg cggccgtcca gagcgggctc cccaccctc 240
 gactcctgcg acccgaccg caccaccacc cgggcccgga ggatgatgaa gctcaagtcg 300
 aaccagaccc gcacctacga cggcgacggc tacaagaagc gggccgcatg cctgtgtttc 360
 cgcagcgaga gcgaggagga ggtgctactc gtgagcagta gtcgccatcc agacagatgg 420

attgtccctg	gaggaggcat	ggagcccag	gaggagccaa	gtgtggcagc	agttcgtgaa	480
gtctgtgagg	aggctggagt	aaaagggaca	ttgggaagat	tagttggaat	ttttgagaac	540
caggagagga	agcacaggac	gtatgtctat	gtgctcattg	tcactgaagt	gctggaagac	600
tgggaagatt	cagttaacat	tggaaggaag	agggaatggt	ttaaaataga	agacgccata	660
aaagtgtctg	agtatcacia	acccgtgcag	gcacatatt	ttgaaacatt	gaggcaaggc	720
tactcagcca	acaatggcac	cccagtcgtg	gccaccacat	actcggtttc	tgctcagagc	780
tcgatgtcag	gcacagatg	actgaagact	tcctgtaaga	gaaatggaaa	ttggaaacta	840
gactgaagtg	caaatcttcc	ctctcaccct	ggctctttcc	acttctcaca	ggcctcctct	900
ttcaaataag	gcattggtgg	cagcaaagaa	agggtgtatt	gataatgttg	ctgtttggtg	960
ttaagtgatg	gggctttttc	ttctgttttt	attgaggggtg	gggggttgggt	gtgtaatttg	1020
taagtacttt	tgtgcatgat	ctgtccctcc	ctcttcccac	ccctgcagtc	ctctgaagag	1080
aggccaacag	ccttcccctg	ccttggattc	tgaagtgttc	ctgtttgtct	tatcctggcc	1140
ctggccagac	gttttctttg	atttttaatt	tttttttttt	attaaaagat	accagtatga	1200
gaaaaaaaaa	aaaaaaaaac	tcgag				1225

<210> 103

<211> 741

<212> DNA

<213> Homo sapien

<400> 103

agaaacctca	atcggattca	gcaaaggaat	ggtgttatta	tcactacata	ccaaatgtta	60
atcaataact	ggcagcaact	ttcaagcttt	agggggccaag	agtttgtgtg	ggactatgtc	120
atcctcgtatg	aagcacataa	aataaaaaacc	tcactacta	agtcagcaat	atgtgctcgt	180
gctattctctg	caagtaatcg	cctcctcctc	acaggaaacc	caatccagaa	taattttacaa	240
gaactatggg	ccctatttga	ttttgcttgt	caaggggtccc	tgctgggaac	attaaaaact	300
tttaagatgg	agtatgaaaa	tcctattact	agagcaagag	agaaggatgc	taccccagga	360
gaaaaagcct	tgggattttaa	aatatctgaa	aacttaatgg	caatcataaa	accctatttt	420
ctcaggagga	ctaaagaaga	cgtacagaag	aaaaagtcaa	gcaaccacaga	ggccagactt	480
aatgaaaaga	atccagatgt	tgatgccatt	tgtgaaatgc	cttccctttc	caggagaaat	540
gatttaatta	tttgatacag	acttgtgcct	ttacaagaag	aaatatacag	gaaattttgtg	600
tcttttagatc	atatcaagga	gttgctaattg	gagacgcgct	cacctttggc	tgagctaggt	660
gtcttaaaga	agctgtgtga	tcacccctag	ctgctgtctg	cacgggcttg	ttgtttgcta	720
aatcttggga	cattctctgc	t				741

<210> 104

<211> 321

<212> DNA

<213> Homo sapien

<400> 104

ttgctctgcg	tcatacaaga	caccaaactg	ctgtgctata	aaagttccaa	ggaccagcag	60
cctcagatgg	aactgccact	ccaaggctgt	aacattacgt	acatcccga	agacagcaaa	120
aagaagaagc	acgagctgaa	gattactcag	cagggcacgg	acccgcttgt	tctcgccgtc	180
cagagcaagg	aacaggccga	gcagtggctg	aaggatgatca	agaagccta	cagtgggtgt	240
agtggccccg	tggattcaga	gtgtcctcct	ccaccaagct	ccccggtgca	caaggcagaa	300
ctggagaaga	aactgtcttc	a				321

<210> 105

<211> 389

<212> DNA

<213> Homo sapien

<400> 105

Met Glu Tyr Lys Gly Glu Leu Ala Ser Tyr Asp Met Arg Leu Arg Arg
 85 90 95
 Lys Leu Asp Leu Phe Ala Asn Val Ile His Val Lys Ser Leu Pro Gly
 100 105 110
 Tyr Met Thr Arg His Asn Asn Leu Asp Leu Val Ile Ile Arg Glu Gln
 115 120 125
 Thr

<210> 116
 <211> 550
 <212> DNA
 <213> Homo sapien

<400> 116
 gaattcggca ccagcctcag agccccccag cccggctacc accccctgcg gaaaggtacc 60
 catctgcatt cctgcccgtc gggacctggt ggacagtcca gcctccttgg cctctagcct 120
 tggctcaccg ctgcctagag ccaaggagct catcctgaat gaccttcccg ccagcactcc 180
 tgcctccaaa tcctgtgact cctccccgcc ccaggacgct tccaccccca ggcccagctc 240
 ggccagtcac ctctgccagc ttgctgccaa gccagcacct tccacggaca gcgtcgccct 300
 gaggagcccc ctgactctgt ccagtccctt caccacgtcc ttcagcctgg gctcccacag 360
 cactctcaac ggagacctct ccgtgcccag ctcttacgtc agcctccacc tgtcccccca 420
 ggtcagcagc tctgtggtgt acggacgctc ccccgatgat gcatttgagt ctcaccccca 480
 tctccgaggg tcatccgtct ctctctccct acccagcatc cctgggggaa agccggccta 540
 ctctttccac 550

<210> 117
 <211> 154
 <212> DNA
 <213> Homo sapien

<400> 117
 ttctgaggga aagccgagt gagtgggcca cccggcggcg gtgacaatga gttttcttgg 60
 aggctttttt ggtcccattt gtgagattga tgttgccctt aatgatgggg aaaccaggaa 120
 aatggcagaa atgaaaactg aggatggcaa agta 154

<210> 118
 <211> 449
 <212> DNA
 <213> Homo sapien

<400> 118
 gaattcggca ccagggcccg cagccccgagt gtgcgcgccca tggcttcgcc gcagctctgc 60
 cgcgcgctgg tgtcggcgca atgggtggcg gaggcgctgc gggccccgcg cgctgggcag 120
 cctctgcagc tgctggacgc ctccctggtac ctgcogaagc tggggcgcca cgcgcgacgc 180
 gagttcgagg agcgccacat cccgggcgcc gctttcttcg acatcgacca gtgcagcgac 240
 cgcacctgc cctacgacca catgctgccc ggggcgcgagc atttcgcgga gtacgcaggc 300
 cgctggggcg tgggcgcggc caccacgtc gtgatctacg acgccagcga ccagggcctc 360
 tactccgcc cgcgctctgt gtggatgttc cgcgccttcg gccaccacgc cgtgtcactg 420
 cttgatggcg gcctccgccca ctggctgcg 449

<210> 119
 <211> 642
 <212> DNA

<213> Homo sapien

<400> 119

gaattcggca	cgagcagtaa	cccgaccgcc	gctgggtcttc	gctggacacc	atgaatcaca	60
ctgtccaaac	cttcttctct	cctgtcaaca	gtggccagcc	ccccaactat	gagatgctca	120
aggaggagca	cgagggtggct	gtgctggggg	cgccccacaa	ccctgctccc	ccgacgtcca	180
ccgtgatcca	catccgcagc	gagacctccg	tgcccgacca	tgctgctctgg	tccctgttca	240
acacctctt	catgaacccc	tgctgcctgg	gcttcatagc	attcgcttac	tccgtgaagt	300
ctagggacag	gaagatggtt	ggcgacgtga	ccggggccca	ggcctatgcc	tccaccgcca	360
agtgcctgaa	catctggggc	ctgattctgg	gcaccccat	gaccattctg	ctcatcgtca	420
tcccagtgt	gatcttcag	gcctatggat	agatcaggag	gcactactga	ggccaggagc	480
tctgcccag	acctgtatcc	cacgtactcc	aacttccatt	cctcgccctg	cccccgagc	540
cgagtctgt	atcagccctt	tatcctcaca	cgcttttcta	caatggcatt	caataaagt	600
cacgtgtttc	tggtgaaaaa	aaaaaaaaa	aaaaaactcg	ag		642

<210> 120

<211> 603

<212> DNA

<213> Homo sapien

<400> 120

gaattcggca	cgagccacaa	cagccactac	gactgcatcc	actggatcca	cggccacccc	60
gtcctccacc	ccgggaacag	ctccccctcc	caaagtgtg	accagcccgg	ccaccacacc	120
catgtccacc	atgtccacaa	tccacacctc	ctctactcca	gagaccaccc	acacctccac	180
agtgtctgacc	accacagcca	ccatgacaag	ggccaccaat	tccacggcca	cacctcctc	240
cactctgggg	acgaccggga	tcctcactga	gtgaccaca	acagccacta	caactgcagc	300
cactggatcc	acggccaccc	tgctcctccac	cccagggaac	acctggatcc	tcacagagcc	360
gagcactata	gccaccgtga	tggtgcccac	cggttccaag	gccaccgcct	cctccactct	420
gggaacagct	cacaccccca	aagtgggtgac	caccatggcc	actatgccc	cagccactgc	480
ctccacgggt	cccagctcgt	ccaccgtggg	gaccacccgc	accctgcag	tgctccccag	540
cagcctgcca	accttcagcg	tgtccactgt	gtcctcctca	gtcctcacca	ccctgagacc	600
cac						603

<210> 121

<211> 178

<212> PRT

<213> Homo sapien

<400> 121

Ser	Glu	Pro	Pro	Ser	Pro	Ala	Thr	Thr	Pro	Cys	Gly	Lys	Val	Pro	Ile
1				5					10					15	
Cys	Ile	Pro	Ala	Arg	Arg	Asp	Leu	Val	Asp	Ser	Pro	Ala	Ser	Leu	Ala
			20					25					30		
Ser	Ser	Leu	Gly	Ser	Pro	Leu	Pro	Arg	Ala	Lys	Glu	Leu	Ile	Leu	Asn
		35				40						45			
Asp	Leu	Pro	Ala	Ser	Thr	Pro	Ala	Ser	Lys	Ser	Cys	Asp	Ser	Ser	Pro
	50					55					60				
Pro	Gln	Asp	Ala	Ser	Thr	Pro	Arg	Pro	Ser	Ser	Ala	Ser	His	Leu	Cys
65					70					75				80	
Gln	Leu	Ala	Ala	Lys	Pro	Ala	Pro	Ser	Thr	Asp	Ser	Val	Ala	Leu	Arg
			85					90						95	
Ser	Pro	Leu	Thr	Leu	Ser	Ser	Pro	Phe	Thr	Thr	Ser	Phe	Ser	Leu	Gly
		100						105					110		
Ser	His	Ser	Thr	Leu	Asn	Gly	Asp	Leu	Ser	Val	Pro	Ser	Ser	Tyr	Val

000230 022950

115 120 125
 Ser Leu His Leu Ser Pro Gln Val Ser Ser Ser Val Val Tyr Gly Arg
 130 135 140
 Ser Pro Val Met Ala Phe Glu Ser His Pro His Leu Arg Gly Ser Ser
 145 150 155 160
 Val Ser Ser Ser Leu Pro Ser Ile Pro Gly Gly Lys Pro Ala Tyr Ser
 165 170 175
 Phe His

<210> 122
 <211> 36
 <212> PRT
 <213> Homo sapien

<400> 122
 Met Ser Phe Leu Gly Gly Phe Phe Gly Pro Ile Cys Glu Ile Asp Val
 1 5 10 15
 Ala Leu Asn Asp Gly Glu Thr Arg Lys Met Ala Glu Met Lys Thr Glu
 20 25 30
 Asp Gly Lys Val
 35

<210> 123
 <211> 136
 <212> PRT
 <213> Homo sapien

<400> 123
 Met Ala Ser Pro Gln Leu Cys Arg Ala Leu Val Ser Ala Gln Trp Val
 1 5 10 15
 Ala Glu Ala Leu Arg Ala Pro Arg Ala Gly Gln Pro Leu Gln Leu Leu
 20 25 30
 Asp Ala Ser Trp Tyr Leu Pro Lys Leu Gly Arg Asp Ala Arg Arg Glu
 35 40 45
 Phe Glu Glu Arg His Ile Pro Gly Ala Ala Phe Phe Asp Ile Asp Gln
 50 55 60
 Cys Ser Asp Arg Thr Ser Pro Tyr Asp His Met Leu Pro Gly Ala Glu
 65 70 75 80
 His Phe Ala Glu Tyr Ala Gly Arg Leu Gly Val Gly Ala Ala Thr His
 85 90 95
 Val Val Ile Tyr Asp Ala Ser Asp Gln Gly Leu Tyr Ser Ala Pro Arg
 100 105 110
 Val Trp Trp Met Phe Arg Ala Phe Gly His His Ala Val Ser Leu Leu
 115 120 125
 Asp Gly Gly Leu Arg His Trp Leu
 130 135

<210> 124
 <211> 133
 <212> PRT
 <213> Homo sapien

<400> 124

000250"02'2950

<212> DNA

<213> Homo sapien

<400> 126

gaattcggca	cgagccaagt	accccctgag	gaatctgcag	cctgcattctg	agtacaccgt	60
atccctcgtg	gccataaagg	gcaaccaaga	gagcccaaaa	gccactggag	tctttaccac	120
actgcagcct	gggagctcta	ttccacctta	caacaccgag	gtgactgaga	ccaccattgt	180
gatcacatgg	acgcctgctc	caagaattgg	ttttaagctg	ggtgtacgac	caagccaggg	240
aggagaggca	ccacgagaag	tgacttcaga	ctcaggaagc	atcgttgtgt	ccggcttgac	300
tccaggagta	gaatacgtct	acaccatcca	agtcctgaga	gatggacagg	aaagagatgc	360
gccaattgta	aacaaagtgg	tgacaccatt	gtctccacca	acaaaacttg	atctggaggc	420
aaaccctgac	actggagtgc	tcacagtctc	ctggagagga	gcaccacccc	agacattact	480
gggtatagaa	ttaccacaac	ccctacaaa				509

<210> 127

<211> 500

<212> DNA

<213> Homo sapien

<400> 127

gaattcggca	cgagccactg	atgtccgggg	agtcagccag	gagcttgggg	aagggaagcg	60
cgcccccg	gcccgtcccg	gagggctcga	tccgcattcta	cagcatgagg	ttctgcccg	120
ttgctgagag	gacgcgtcta	gtcctgaagg	ccaaggggaat	caggcatgaa	gtcatcaata	180
tcaacctgaa	aaataagcct	gagtggttct	ttaagaaaaa	tccctttgg	ctgggtgccag	240
ttctggaaaa	cagtcagggt	cagctgatct	acgagtctgc	catcacctgt	gagtacctgg	300
atgaagcata	cccaggggaag	aagctggttc	cggatgaccc	ctatgagaaa	gcttgccaga	360
agatgatctt	agagttggtt	tctaagggtc	catccttgg	aggaagcttt	attagaagcc	420
aaaataaaga	agactatgct	ggcctaaaag	aagaatttcg	taaagaattt	accaagctag	480
aggaggttct	gactaataag					500

<210> 128

<211> 500

<212> DNA

<213> Homo sapien

<400> 128

agctttctc	tgctgccgct	cggtcacgct	tgtgcccgaa	ggaggaaaca	gtgacagacc	60
tggagactgc	agttctctat	ccttcacaca	gctctttcac	catgcctgga	tcaattcctt	120
tgaatgcaga	agcttgctgg	ccaaaagatg	tgggaattgt	tgcccttgag	atctattttc	180
cttctcaata	tggtgatcaa	gcagagttgg	aaaaatatga	tggtgtagat	gctggaaagt	240
ataccattgg	cttggggccag	gccaaagatg	gcttctgcac	agatagagaa	gatattaact	300
ctctttgcat	gactgtggtt	cagaatctta	tggagagaaa	taacctttcc	tatgattgca	360
ttgggcggct	ggaagttgga	acagagacaa	tcacgacaaa	atcaaagtct	gtgaagacta	420
atttgatgca	gctgtttgaa	gagtctggga	atacagatat	agaaggaatc	gacacaacta	480
atgcatgcta	tggaggcaca					500

<210> 129

<211> 497

<212> DNA

<213> Homo sapien

<400> 129

gaattcggca	cgagcagagg	tctccagagc	cttctctctc	ctgtgcaaaa	tggcaactct	60
taaggaaaaa	ctcattgcac	cagttgcgga	agaagaggca	acagttccaa	acaataagat	120

cactgtagtg	ggtgttgac	aagttggtat	ggcgtgtgct	atcagcattc	tgggaaagtc	180
tctggctgat	gaacttgctc	ttgtggatgt	tttggaagat	aagcttaaag	gagaaatgat	240
ggatctgcag	catgggagct	tatttcttca	gacacctaaa	attgtggcag	ataaagatta	300
ttctgtgacc	gccaatctta	agattgtagt	ggtaactgca	ggagtccgtc	agcaagaagg	360
ggagagtcgg	ctcaatctgg	tgcagagaaa	tgtaaatgtc	ttcaaattca	ttattcctca	420
gatcgtcaag	tacagtcctg	attgcatcat	aattgtgggt	tccaaccag	tggacattct	480
tacgtatggt	acctgga					497

<210> 130

<211> 383

<212> DNA

<213> Homo sapien

<400> 130

gaattcggca	cgagggccgc	ggctgccgac	tgggtcccct	gccgctgtcg	ccaccatggc	60
tccgcaccgc	cccgcgccc	cgctgctttg	cgcgctgtcc	ctggcgctgt	gcgcgctgtc	120
gctgcccgtc	cgcgcgccca	ctgcgtcgcg	gggggcgctc	caggcggggg	cgccccaggg	180
gcgggtgccc	gaggcgcggc	ccaacagcat	ggtggtggaa	caccccgagt	tcctcaaggc	240
agggaaggag	cctggcctgc	agatctggcg	tgtggagaaa	gttcgatctg	gtggcccgtg	300
cccaccaacc	tttatggaga	cttcttcacg	ggcgacgcct	acgtcatcct	gaagacagtg	360
cagcttaaga	acggaaaatc	ttg				383

<210> 131

<211> 509

<212> DNA

<213> Homo sapien

<400> 131

gaattcggca	cgagagtcag	ccgcattctc	ttttgcgtcg	ccagccgagc	cacatcgctc	60
agacaccatg	gggaagggtga	aggtcggagt	caacggattt	ggtcgtattg	ggcgccctgg	120
caccagggct	gcttttaact	ctggtaaagt	ggatattgtt	gccatcaatg	accccttcat	180
tgacctcaac	tacatggttt	acatgttcca	atatgattcc	acccatggca	aattccatgg	240
caccgtcaag	gctgagaacg	ggaagcttgt	catcaatgga	aatcccatca	ccatcttcca	300
ggagcgagat	ccctccaaaa	tcaagtgggg	cgatgctggc	gctgagtacg	tcgtggagtc	360
cactggccgt	cttcaccacc	atggagaagg	ctggggctca	tttgcagggg	ggagccaaaa	420
gggtcatcat	ctctgcccc	tctgctgacg	cccccatgtt	cgtcatgggt	gtgaaccatg	480
agaagtatga	caacagcctc	aagatcatc				509

<210> 132

<211> 357

<212> DNA

<213> Homo sapien

<400> 132

gaattcggca	cgagtaagaa	gaagccccta	gaccacagct	ccacaccatg	gactggacct	60
ggaggatcct	cttcttggtg	gcagcagcaa	caggtgccca	ctcccagggtg	caactgggtg	120
aatctgggtc	tgagttgaag	aagcctgggg	cctcagtga	ggtttctctg	aaggcttctg	180
gacacatctt	cagtatctat	ggtttgaatt	gggtgcgaca	ggcccctggt	caaggccttg	240
agtggatggg	atggatcaaa	gtcgacactg	cgaacccaac	gtatgccag	ggcttcacag	300
gacgatttgt	cttctccctg	gacacctctg	tcagcacggc	atatctgcag	atcagca	357

<210> 133

<211> 468

<212> DNA

attcgttgta	tgaatcttat	tacaagcagg	tccatccggc	atacacaggg	aggggtggggg	120
cgagtgaagc	tgcgcttttt	ctaaagaagt	ctggcctctc	ggacattatc	cttgggaaga	180
tatgggactt	ggccgatcca	gaaggtaaag	ggttcttgga	caaacagggg	ttctatgttg	240
cactgagact	ggtggcctgt	gcacagagtg	gccatgaagt	taccttgagc	aatctgaatt	300
tgagcatgcc	accgcctaaa	tttcacgaca	ccagcagccc	tctgatggtc	acaccgccct	360
ctgcagaggc	ccactgggct	gtgaggggtg	aagaaaaggc	caaatttgat	gggatttttg	420
aaagcctctt	gcccataaat	ggtttgcctc	ctggagacaa	agtcaagcca	gtcctcatga	480
actcaaagct	gcctcttgat	gtc				503

<210> 149

<211> 1061

<212> DNA

<213> Homo sapien

<400> 149

gaattcggca	cgaggccttt	tccagcaacc	ccaaggtcca	ggtggaggcc	atcgaagggg	60
gagccctgca	gaagctgctg	gtcatcctgg	ccacggagca	gccgctcact	gcaaagaaga	120
aggtcctgtt	tgcactgtgc	tccctgctgc	gccacttccc	ctatgccag	cggcagttcc	180
tgaagctcgg	ggggctgcag	gtcctgagga	ccctggtgca	ggagaagggc	acggaggtgc	240
tgcgcgtgcg	cgtggtcaca	ctgctctacg	acctggtcac	ggagaagatg	ttcgccgagg	300
aggaggctga	gctgaccag	gagatgtccc	cagagaagct	gcagcagtat	cgccaggtac	360
acctcctgcc	aggcctgtgg	gaacagggct	ggtgctgagat	cacggcccac	ctcctggcgc	420
tgcccagagca	tgatgcccg	gagaagggtg	tgcagacact	gggcgtcctc	ctgaccacct	480
gccgggaccg	ctaccgtcag	gacccccagc	tccgcaggac	actggccagc	ctgcaggctg	540
agtaccaggt	gctggccagc	ctggagctgc	aggatggtga	ggacgagggc	tacttccagg	600
agctgctggg	ctctgtcaac	agcttgctga	aggagctgag	atgaggcccc	acaccagtac	660
tggactggga	tgcgctagt	gaggtgagg	ggtgccagcg	tgggtgggct	tctcaggcag	720
gaggacatct	tggcagtgt	ggcttgcca	ttaaatggaa	acctgaaggc	catcctcttt	780
ctgctgtgtg	tctgtgtaga	ctgggcacag	ccctgtggcc	gggggggtcag	gtgagtgggt	840
gggtgatggg	ctctgctgac	gtgcagggct	cagcccaggg	catccaggaa	caggctccag	900
ggcaggaacc	tgggccag	agttgcaagt	ctctgcttct	taccaagcag	cagctctgta	960
ccttggggaag	tgccttaatt	gctctgagct	tgtttcctca	tctgtcagga	gtgccattaa	1020
aggagaaaaa	tcacgtaaaa	aaaaaaaaaa	aaaaactoga	g		1061

<210> 150

<211> 781

<212> DNA

<213> Homo sapien

<400> 150

gaattcggca	cgagaaatgg	cggcaggggt	cgaagcggca	gccgaagtgg	cggcgacaga	60
acccaaaatg	gaggaagaga	gcggcgcgcc	ctgcgtgccg	agcggcaacg	gagctccggg	120
cccgaagggt	gaagaacgac	ctactcagaa	tgagaagagg	aaggagaaaa	acataaaaag	180
aggaggcaat	cgctttgagc	catattccaa	cccaactaaa	agatacagag	ccttcattac	240
aaatatacct	tttgatgtga	aatggcagtc	acttaaagac	ctggttaaag	aaaaagttgg	300
tgaggtaaca	tacgtggagc	tcttaatgga	cgctgaagga	aagtcaaggg	gatgtgctgt	360
tgttgaaatc	aagatggagg	agagcatgaa	aaaagctgct	gaagttctaa	acaagcatag	420
tctgagtggg	aggccactga	aagtcaagga	agatcctgat	ggtgaacatg	caaggagagc	480
aatgcaaaag	gctggaagac	ttggaagcac	agtatttgta	gcaaactctg	attataaagt	540
tggctggaag	aaactgaagg	aagtatttag	tatggctggg	gtgggtggcc	gagcagacat	600
tctggaagat	aaagatggga	aaagtcgtgg	aataggcatt	gtgacttttg	aacagtcctat	660
tgaagctgtg	caagcaatat	ctatgtttta	tggccagttg	ctgtttgata	gaccgatgca	720
cgtcaagatg	gatgagaggg	ctttaccaa	gggagacttt	tttcctcctg	aacgccacag	780
c						781

<210> 151
 <211> 3275
 <212> DNA
 <213> Homo sapien

<400> 151

cttaagtgga	tcctgcatca	ggagggagca	gacaccggag	aaagaaaaac	aagttgtgct	60
gtttgaggaa	gcaagttgga	cctgcactcc	agcctgtgga	gatgaacct	ggactgtgat	120
tctgctatcc	agtatgttgg	ctgaccacag	gctcaaaactg	gaggattata	aggatcgcc	180
gaaaagtgga	gagcatctta	atccagacca	gttggaagct	gtagagaaat	atgaagaagt	240
gctacataat	ttggaatttg	ccaaggagct	tcaaaaaacc	ttttctgggt	tgagcctaga	300
tctactaaaa	gcgcaaaaga	aggcccagag	aaggggagcac	atgctaaaac	ttgaggctga	360
gaagaaaaag	cttcgaacta	tacttcaagt	tcagtatgta	ttgcagaact	tgacacagga	420
gcacgtacaa	aaagacttca	aaggggggtt	gaatggtgca	gtgtatttgc	cttcaaaaga	480
acttgactac	ctcatthaagt	tttcaaaact	gacctgccct	gaaagaaatg	aaagtctgag	540
acaaacactt	gaaggatcta	ctgtctaaat	tgctgaactc	aggctatttt	gaaagtatcc	600
cagttcccaa	aaatgccaa	gaaaaggaag	taccactgga	ggaagaaatg	ctaatacaat	660
cagagaaaaa	aacacaatta	tcgaagactg	aatctgtcaa	agagtcagag	tctctaattg	720
aatttgccca	gccagagata	caaccacaag	agtttcttaa	cagacgctat	atgacagaag	780
tagattattc	aaacaaacaa	ggcgaagagc	aaccttgga	agcagattat	gctagaaaac	840
caaactctcc	aaaacgttgg	gatatgctta	ctgaaccaga	tggtcaagag	aagaaacagg	900
agtcctttaa	gtcctgggag	gcttctggta	agcaccagga	ggtatccaag	cctgcagttt	960
ccttagaaca	gaggaaacaa	gacacctcaa	aactcaggtc	tactctgccc	gaagagcaga	1020
agaagcagga	gatctccaaa	tccaagccat	ctcctagcca	gtggaagcaa	gatacaccta	1080
aatccaaagc	agggtatggt	caagaggaac	aaaagaaaca	ggagacacca	aagctgtggc	1140
cagttcagct	gcagaaagaa	caagatccaa	agaagcaaac	tccaaagtct	tggacacctt	1200
ccatgcagag	cgaacagaac	accaccaagt	catggaccac	tcccatgtgt	gaagaacagg	1260
attcaaaaaca	gccagagact	ccaaaatcct	gggaaaacaa	tgttgagagt	caaaaacact	1320
ctttaacatc	acagtcacag	atttctccaa	agtcctgggg	agtagctaca	gcaagcctca	1380
taccaaatga	ccagctgctg	cccaggaagt	tgaacacaga	acccaaagat	gtgcctaagc	1440
ctgtgcatca	gcctgtaggt	tcttctctta	cccttccgaa	ggatccagta	ttgaggaaag	1500
aaaaactgca	ggatctgatg	actcagattc	aaggaacttg	taactttatg	caagagtctg	1560
ttcttgactt	tgacaaacct	tcaagtgcaa	ttccaacgtc	acaaccgcct	tcagctactc	1620
caggtagccc	cgtagcatct	aaagaacaaa	atctgtccag	tcaaagtgat	tttcttcaag	1680
agccgttaca	ggtattttaa	gttaatgcac	ctctgcctcc	acgaaaagaa	caagaaataa	1740
aagaatcccc	ttattcacct	ggctacaatc	aaagttttac	cacagcaagt	acacaaacac	1800
caccccagtg	ccaactgcc	tctatacatg	tagaacaac	tgtccattct	caagagactg	1860
cagcaaatta	tcactctgat	ggaactattc	aagtaagcaa	tggtagcctt	gccttttacc	1920
cagcacagac	gaatgtgttt	cccagaccta	ctcagccatt	tgtcaatagc	cggggatctg	1980
ttagaggatg	tactcgtgg	gggagattaa	taaccaattc	ctatcggtcc	cctgggtggt	2040
ataaagggtt	tgatacttat	agaggactcc	cttcaatttc	caatggaaat	tatagccagc	2100
tgcagttcca	agctagagag	tattctggag	caccttattc	ccaaagggat	aatttccagc	2160
agtgttataa	gcgaggagg	acatctggtg	gtccacgagc	aaattcgaga	gcagggtgga	2220
gtgattcttc	tcaggtgagc	agcccagaaa	gagacaacga	aacctttaac	agtgggtgact	2280
ctggacaagg	agactcccg	agcatgaccc	ctgtggatgt	gccagtgaca	aatccagcag	2340
ccaccatact	gccagtacac	gtctaccctc	tgctcagca	gatgcgagtt	gccttctcag	2400
cagccagaac	ctctaactctg	gcccctggaa	ctttagacca	acctattgtg	tttgatcttc	2460
ttctgaacaa	cttaggagaa	acttttgatc	ttcagcttgg	tagattttaa	tgcccagtga	2520
atggcactta	cgttttcatt	tttcacatgc	taaagctggc	agtgaatgtg	ccactgtatg	2580
tcaacctcat	gaagaatgaa	gaggctcttg	tatcagccta	tgccaatgat	ggtgctccag	2640
accatgaaac	tgctagcaat	catgcaattc	ttcagctctt	ccaggggagac	cagatatggt	2700
tacgtctgca	caggggagca	atttatggaa	gtagctggaa	atattctacg	ttttcaggct	2760
atcttcttta	tcaagattga	aagtcagtac	agtattgaca	ataaaaaggat	ggtgttctaa	2820

<211> 2109
 <212> DNA
 <213> Homo sapien

<400> 153

cagagagccc	caggcatcga	ggagaaggcg	goggagaatg	gggccctggg	gtcccccgag	60
agagaagaga	aagtgtctgga	gaatggggag	ctgacacccc	caaggaggga	ggagaaagcg	120
ctggagaatg	gggagctgag	gtccccagag	gccggggaga	aggtgtctgt	gaatgggggc	180
ctgacacccc	caaagagcga	ggacaagggt	tcagagaatg	ggggcctgag	attccccagg	240
aacacggaga	ggccaccaga	gactgggcct	tggagagccc	cagggccctg	ggagaagacg	300
cccgagagtt	gggtccagc	ccccacgatc	ggggagccag	ccccagagac	ctctctggag	360
agagcccctg	caccacagcg	agtgtgtctc	tcccggaaacg	gcggggagac	agcccctggc	420
ccccttgggc	cagcccccaa	gaacgggacg	ctggaacccg	ggaccgagag	gagagccccc	480
gagactgggg	gggcgccgag	agccccaggg	gctgggaggc	tggacctcgg	gagtgggggc	540
cgagccccag	tgggcacggg	gacggccccc	ggcggcgggc	ccggaagcgg	cgtggacgca	600
aaggccggat	gggtagacaa	cacgaggccg	cagccaccgc	cgccaccgct	gccaccgcca	660
ccggaggcac	agccgaggag	gctggagcca	gcgccccga	gagccaggcc	ggaggtggcc	720
cccaggggag	agcccggggc	cccagacagc	agggccggcg	gagacacggc	actcagcggg	780
gacggggacc	ccccaaagcc	cgagaggaag	ggccccgaga	tgccacgact	attcttggac	840
ttgggacccc	ctcaggggaa	cagcgagcag	atcaaagcca	ggctctcccg	gctctcgtctg	900
gcgctgccgc	cgctcacgct	cacgccattc	ccggggccgg	gcccgcggcg	gccccctgg	960
gagggcgcg	acgccggggc	ggctggcggg	gagggccggc	gggcgggagc	gccggggccg	1020
gcggaggagg	acggggagga	cgaggacgag	gacgaggagg	aggacgagga	ggcggcgggc	1080
ccgggcgcgg	cgccggggcc	gcggggcccc	gggagggcgc	gagcagcccc	ggtgcccgctc	1140
gtggtgagca	gcgccgacgc	ggacgcggcc	cgcccgcctgc	gggggctgct	caagtctccg	1200
cgccggggccg	acgagccaga	ggacagcgag	ctggagagga	agcgcaagat	ggtctccttc	1260
cacggggacg	tgacctcta	cctcttcgac	caggagacgc	caaccaacga	gctgagcgctc	1320
caggcccccc	ccgaggggga	cacggacccg	tcaacgcctc	cagcgcccc	gacacctccc	1380
caccccgcca	cccccgagga	tgggtttccc	agcaacgaca	gcggcttttg	aggcagtttc	1440
gagtgggcgg	aggatttccc	cctcctcccc	cctccaggcc	ccccgctgtg	cttctcccgc	1500
ttctccgtct	cgctgcgct	ggagaccccc	gggccacccg	cccgggcccc	cgacgcccgg	1560
cccgaggcc	ccgtggagaa	ttgattcccc	gaagacccga	ccccgctgca	ccctcagaag	1620
aggggttag	aatggaatcc	tctgtggatg	acggcgccac	tgccaccacc	gcagacgccc	1680
cctctgggga	ggcccccgag	gctgggcccct	ccccctccca	ctcccctacc	atgtgccaaa	1740
cgggaggccc	cgggcccccg	cccccccagc	ccccagatg	gtccccctga	ccccctgac	1800
cccctcggag	ccaaatgagg	caggaatccc	ccgcgccctc	catagagagc	cgctttctc	1860
ggaactgaac	tgaactcttt	tgggcctgga	gcccctcgac	acagcggagg	tcctcctca	1920
cccactcctg	gcccagaca	ggggcccgag	gcttcgggga	cccgacccc	ccatttcgcg	1980
tctccccttt	ccctccccag	cccggcccct	ggaggggcct	ctggttcaaa	ccttcgctg	2040
gcattttcac	attattttaa	aaagacaaaa	acaacttttt	ggaggaaaaa	aaaaaaaaaa	2100
aaactcgag						2109

<210> 154
 <211> 1411
 <212> DNA
 <213> Homo sapien

<400> 154

gaattcggca	ccaggggaga	tgaggaagtt	cgatgttcct	agcatggagt	ctacccttaa	60
ccagccagcc	atgctagaga	cgttatactc	agatccacat	taccgagccc	atttcccaa	120
cccaagacct	gatacaata	aggatgtata	caaagtattg	ccagaatcca	agaaggcacc	180
gggcagtgg	gcagtattttg	agaggaacgg	accacatgct	agcagtagtg	gggtgctccc	240
tttgggactc	cagcctgcgc	ctggactttc	caagtcacta	tcctctcagg	tgtggcaacc	300
aagtccctgac	ccttggcatc	ctggagaaca	atcctgtgaa	ctcagtactt	gtcgacagca	360

tgggaagagg ccgagcgtgg ctttatcttg cactcatgca aaagaaactg gcagattatc 660
 tgaaagtgtc tatagacaat aaacatctct taagcgagtt ctatgagcct gaggctttaa 720
 tgatggagga agaagggatg gtgattgttg gtctgctggt gggactcaat gttctcgatg 780
 ccaatctctg cttgaaagga gaagacttgg attctcaggt tggagtaata gatttttccc 840
 tctaccttaa ggatgtgcag gatcttgatg gtggcaagga gcatgaaaga attactgatg 900
 tccttgatca aaaaaattat gtggaagaac ttaaccggca cttgagctgc acagttgggg 960
 atcttcaaac caagatagat ggcttggaaa agactaactc aaagcttcaa gaagagcttt 1020
 cagctgcaac agaccgaatt tgctcacttc aagaagaaca gcagcagtta agagaacaaa 1080
 atgaattaat tcgagaaaga agtgaaaaga gtgtagagat aacaaaacag gataccaaag 1140
 ttgagctgga gacttacaag caaactcggc aaggctctgga tgaaatgtac agtgatgtgt 1200
 ggaagcagct aaaagaggag aagaaagtcc ggttggaact ggaaaaagaa ctggagttac 1260
 aaattggaat gaaaaccgaa atggaaattg caatgaagtt actggaaaag gacacccacg 1320
 agaagcagga cacactagtt gccctccgcc agcagctgga agaagtcaaa gcgattaatt 1380
 tacagatgtt tcacaaagct cagaatgcag agagcagttt gcagcagaag aatgaagcca 1440
 tcacatcctt tgaaggaaaa accaaccaag ttatgtccag catgaaacaa atggaagaaa 1500
 gggtgcagca ctcgagcgcg gcgaggcagg gggctgagga gcgagccac aagctgcagc 1560
 aggagctggg cgggaggatc ggcgccctgc agctgcagct ctcccagctg cacgagcaat 1620
 gctcaagcct ggagaaagaa ttgaaatcag aaaaagagca aagacagget cttcagcgcg 1680
 aattacagca cgagaaagac acttcctctc tactcaggat ggagctgcaa caagtggaaag 1740
 gactgaaaaa ggagttgcgg gagcttcagg acgagaaggc agagctgcag aagatctgcg 1800
 aggagcagga acaagccctc caggaaatgg gcctgcacct cagccagtc aagctgaaga 1860
 tggagatat aaaagaagtg aaccaggcac tgaagggcca cgctggctg aaagatgacg 1920
 aagcgacaca ctgtaggcag tgtgagaagg agttctccat ttcccggaga aagcaccact 1980
 gccggaactg tggccacatc ttctgcaaca cctgctccag caacgagctg gccctgccct 2040
 cctaccccaa gccggtgcga gtgtgcgaca cctgccacac cctgctcctg cagcgctgct 2100
 cctccacggc ctctctgaacg tccgtcctca ggagcacagc ctacgggaca gtgccaaacc 2160
 ctgtgggtct ccaggggctt gggaaatgtg ttctttccca agagtatcaa aggaaagaat 2220
 caaattttct gcccggctac tggcactcca gaagacagcg tgccggaacc ggcagctctc 2280
 acctttctgt gacttggttcg gaattaactc ctctggatgg aaacttccat cttacttggt 2340
 tacatcacgg ctctggttca gatacaactt catgattttg ctactatcat ttttcacttt 2400
 tcaaagaatt taacctatct tacagcagtt cagttctgct agtgagtagt tttcctctcc 2460
 taccttctct ctaaaaacct gattcatgca cagcgtttga cacacatgga gtctgccagt 2520
 gtgccttctc tgcttcagac aagagatctg ccatttcatg cccttggtgac tacctatcat 2580
 tggccctgca ataaaatcat ttatttttca aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2640
 aaaaaaaaaa aaaaaaaaaa aactcgag 2668

<210> 157

<211> 2313

<212> DNA

<213> Homo sapien

<400> 157

gaattcggca ccaggccggg cgggcgcctc agccatggcc ctgcgcaagg aactgctcaa 60
 gtccatctgg tacgccttta ccgcgtgga cgtggagaag agtggcaaag tctccaagtc 120
 ccagctcaag gtgctgtccc acaacctgta cacggctctg cacatcccc atgaccccg 180
 ggcctggag gaacacttcc gagatgatga tgacggccct gtgtccagcc agggatacat 240
 gccctacctc aacaagtaca tccctggaca ggtggaggag ggggcttttg ttaaagagca 300
 ctttgatgag ctgtgctgga cgctgacggc caagaagaac tatcgggcag atagcaacgg 360
 gaacagtatg ctctccaatc aggatgcctt ccgcctctgg tgccctctca acttcctgtc 420
 tgaggacaag taccctctga tcatggttcc tgatgaggtg gaatacctgc tgaaaaaggt 480
 actcagcagc atgagcttgg aggtgagctt ggggtgagctg gaggagcttc tggcccagga 540
 ggcccagggt gccagacca ccggggggct cagcgtctgg cagttcctgg agctcttcaa 600
 ttccggccgc tgcttgccgg gcgtgggccc ggacaccctc agcatggcca tccacgaggt 660
 ctaccaggag ctcatccaag atgtcctgaa gcagggtac ctgtggaagc gagggcacct 720

gagaaggaac	tgggccgaac	gctgggttcca	gctgcagccc	agctgcctct	gctactttgg	780
gagtgaagag	tgcaaaagaga	aaaggggcat	tatcccgcctg	gatgcacact	gctgcgtgga	840
ggtgctgcc	gaccgcgacg	gaaagcgctg	catgtttctgt	gtgaagacag	ccacccgcac	900
gtatgagatg	agcgctcag	acacgcgcca	gcgccaggag	tggacagctg	ccatccagat	960
ggcgatccgg	ctgcaggccg	aggggaagac	gtccctacac	aaggacctga	agcagaaacg	1020
gcgcgagcag	cgggagcagc	gggagcggcg	ccggggcgcc	aaggaagagg	agctgctgcg	1080
gctgcagcag	ctgcaggagg	agaaggagcg	gaagctgcag	gagctggagc	tgctgcagga	1140
ggcgagcgg	caggccgagc	ggctgctgca	ggaggaggag	gaacggcgcc	gcagccagca	1200
ccgcgagctg	cagcaggcgc	tcgagggcca	actgcgcgag	gcggagcagg	cccgggcctc	1260
catgcaggct	gagatggagc	tgaaggaggga	ggaggctgcc	cggcagcggc	agcgcatcaa	1320
ggagctggag	gagatgcagc	agcggttgca	ggaggccctg	caactagagg	tgaaagctcg	1380
gcgagatgaa	gaatctgtgc	gaatcgctca	gaccagaactg	ctggaagagg	aggaagagaa	1440
gctgaagcag	ttgatgcagc	tgaaggaggga	gcaggagcgc	tacatcgaac	gggcgcagca	1500
ggagaaggaa	gagctgcagc	aggagatggc	acagcagagc	cgctccctgc	agcaggccca	1560
gcagcagctg	gaggaggtgc	ggcagaaccg	gcagagggct	gacgaggatg	tggaggctgc	1620
ccagagaaaa	ctgcgccagg	ccagcaccaa	cgtgaaacac	tggaatgtcc	agatgaaccg	1680
gctgatgcat	ccaattgagc	ctggagataa	gcgtccggtc	acaagcagct	ccttctcagg	1740
cttccagccc	cctctgcttg	cccaccgtga	ctcctcccta	aagcgctga	cccgtctggg	1800
atcccagggc	aacaggaccc	cctcgcccaa	cagcaatgag	cagcagaagt	ccctcaatgg	1860
tggggatgag	gctcctgccc	cggcttccac	ccctcaggaa	gataaaactg	atccagcacc	1920
agaaaattag	cctctcttag	ccccttgttc	ttcccaatgt	catatccacc	aggacctggc	1980
cacagctggc	ctgtgggtga	tcccagctct	tactaggaga	gggagctgag	gtcctggtgc	2040
caggggccc	ggccctccaa	ccataaacag	tccaggatgg	aacctggttc	acccttcata	2100
ccagctccaa	gccccagacc	atgggagctg	tctgggatgt	tgatccttga	gaacttggcc	2160
ctgtgcttta	gacccaagga	cccgatccct	gggctaggaa	agagagaaca	agcaagccgg	2220
ggctacctgc	ccccaggtgg	ccaccaagtt	gtggaagcac	atttctaaat	aaaaactgct	2280
cttagaatga	aaaaaaaaaa	aaaaaaactc	gag			2313

<210> 158

<211> 2114

<212> DNA

<213> Homo sapien

<400> 158

gaattcggca	cgaggaagaa	ctcgctctctg	ttgagtgtaa	gtagccaaac	aataaccaag	60
gagaataaca	gaaatgtcca	tttgagcac	tcagagcaga	atcctggttc	atcagcaggt	120
gacacctcag	cagcgcacca	ggtggtttta	ggagaaaact	tgatagccac	agccctttgt	180
ctttctggca	gtgggtctca	gtctgatttg	aaggatgtgg	ccagcacagc	aggagaggag	240
ggggacacaa	gccttcggga	gagcctccat	ccagtcactc	ggtctcttaa	ggcaggggtgc	300
catactaagc	agcttgccctc	caggaattgc	tctgaagaga	aatccccaca	aacctccatc	360
ctaaaggaag	gtaacaggga	cacaagcttg	gatttccgac	ctgtagtgtc	tccagcaaat	420
ggggttgaag	gagtcagagt	ggatcaggat	gatgatcaag	atagctcttc	cctgaagctt	480
tctcagaaca	ttgctgtaca	gactgacttt	aagacagctg	attcagaggt	aaacacagat	540
caagatattg	aaaagaattt	ggataaaatg	atgacagaga	gaacctgtgt	gaaagagcgt	600
taccaggagg	tcctggacaa	acagaggcaa	gtggagaatc	agctccaagt	gcaattaaag	660
cagcttcagc	aaaggagaga	agaggaaatg	aagaatcacc	aggagatatt	aaaggctatt	720
caggatgtga	caataaagcg	ggaagaaaca	aagaagaaga	tagagaaaga	gaagaaggag	780
tttttgcaga	aggagcagga	tctgaaagct	gaaattgaga	agctttgtga	gaagggcaga	840
agagaggtgt	gggaaatgga	actggataga	ctcaagaatc	aggatggcga	aataaatagg	900
aacattatgg	aagagactga	acgggcctgg	aaggcagaga	tcttatcact	agagagccgg	960
aaagagttac	tggtactgaa	actagaagaa	gcagaaaaag	aggcagaatt	gcaccttact	1020
tacctcaagt	caactcccc	aacactggag	acagttcggt	ccaaacagga	gtgggagacg	1080
agactgaatg	gagttcggat	aatgaaaaag	aatgttcgtg	accaatttaa	tagtcatatc	1140
cagttagtga	ggaacggagc	caagctgagc	agccttccctc	aaatccctac	tcccacttta	1200

acc

543

<210> 174
 <211> 548
 <212> DNA
 <213> Homo sapien

<400> 174

gaattcggca	cgagaaatgg	cggcaggggt	cgaagcggcg	gcggaggtgg	cggcgacgga	60
gatcaaaaatg	gaggaagaga	gcggcgcgcc	cggcgtgccc	agcggcaacg	gggctccggg	120
ccctaagggt	gaaggagaac	gacctgctca	gaatgagaag	aggaaggaga	aaaacataaa	180
aagaggaggc	aatcgctttg	agccatatgc	caatccaact	aaaagataca	gagccttcat	240
tacaaacata	ccttttgatg	tgaaatggca	gtcacttaaa	gacctgggta	aagaaaaagt	300
tggtgaggta	acatacgtgg	agctcttaat	ggacgctgaa	ggaaagtcaa	ggggatgtgc	360
tgttgttgaa	ttcaagatgg	aagagagcat	gaaaaaagct	gcggaagtcc	taaacaagca	420
tagtctgagc	ggaagaccac	tgaaagtcaa	agaagatcct	gatgggtgaac	atgccaggag	480
agcaatgcaa	aaggtgatgg	ctacgactgg	tgggatgggt	atgggaccag	gtggcccagg	540
aatgatta						548

<210> 175
 <211> 604
 <212> DNA
 <213> Homo sapien

<400> 175

gaattcggca	ccagaggacc	tccaggacat	gttcatcgtc	cataccatcg	aggagattga	60
gggcctgatc	tcagcccatg	accagttcaa	gtccaccctg	ccggacgccg	atagggagcg	120
cgaggccatc	ctggcccatcc	acaaggaggc	ccagaggatc	gctgagagca	accacatcaa	180
gctgtcgggc	agcaacccct	acaccaccgt	caccccgcaa	atcatcaact	ccaagtggga	240
gaaggtgcag	cagctgggtgc	caaaacggga	ccatgccctc	ctggaggagc	agagcaagca	300
gcagtccaac	gagcacctgc	gccgccagtt	cgccagccag	gccaatgttg	tggggccctg	360
gatccagacc	aagatggagg	agatcgggcg	catctccatt	gagatgaacg	ggaccctgga	420
ggaccagctg	agccacctga	agcagtatga	acgcagcatc	gtggactaca	agcccaacct	480
ggacctgctg	gagcagcagc	accagcttat	ccaggaggcc	ctcatcttcg	acaacaagca	540
caccaactat	accatggagc	acatccgcgt	gggctgggag	cagctgctca	ccaccattgc	600
ccgg						604

<210> 176
 <211> 486
 <212> DNA
 <213> Homo sapien

<400> 176

gaattcggca	ccagccaagc	tcaactattga	atccacgcgc	ttcaatgtcg	cagaggggaa	60
ggaggttctt	ctactcgccc	acaacctgcc	ccagaatcgt	attgggttaca	gctgggtacaa	120
aggcgaaaga	gtggatggca	acagtctaat	tgtaggatat	gtaataggaa	ctcaacaagc	180
taccccaggg	cccgcataca	gtggctcgaga	gacaatatac	cccaatgcat	ccctgctgat	240
ccagaacgtc	accagaatg	acacaggatt	ctatacccta	caagtcataa	agtcagatct	300
tgtgaatgaa	gaagcaaccg	gacagttcca	tgtatacccg	gagctgcccc	agccctccat	360
ctccagcaac	aactccaacc	ccgtggagga	caaggatgct	gtggccttca	cctgtgaacc	420
tgaggttcag	aacacaacct	acctgtgggtg	ggtaaattggt	cagagcctcc	cggtcagtc	480
caaggc						486

<210> 177

000260"02'29360

<211> 387
 <212> DNA
 <213> Homo sapien

<400> 177
 gaattcggca ccaggggacag cagaccagac agtcacagca gccttgacaa aacgttcctg 60
 gaactcaagc tcttctccac agaggaggac agagcagaca gcagagacca tggagtctcc 120
 ctcggccctt cccacagat ggtgcatccc ctggcagagg ctctgtctca cagcctcact 180
 tctaaccctt tggaacccgc ccaccaactgc caagctcact attgaatcca cgccgttcaa 240
 tgtcgcagag ggggaaggagg tgtttctact tgtccacaat ctgccccagc atcttttttg 300
 ctacagctgg tacaaggtg aaagagtggg tggcaaccgt caaattatag gatattgtaat 360
 aggaactcaa caagctaccc cagggcc 387

<210> 178
 <211> 440
 <212> DNA
 <213> Homo sapien

<400> 178
 gaattcggca cgaggagaag cagaaaaaca aggaatttag ccagacttta gaaaatgaga 60
 aaaatacctt actgagtcag atatcaacaa aggatgggtga actaaaaatg cttcaggagg 120
 aagtaaccaa aatgaacctg tttaatcagc aaatccaaga agaactctct agagttacca 180
 aactaaagga gacagcagaa gaagagaaaag atgatttgga agagaggctt atgaatcaat 240
 tagcagaact taatggaagc attgggaatt actgtcagga tgttacagat gcccaaataa 300
 aaaatgagct attggaatct gaaatgaaga accttaaaaa gtgtgtgagt gaattggaag 360
 aagaaaagca gcagttagtc aaggaaaaaa ctaagggtgga atcagaaata cgaaaggaaat 420
 atttgagaa aatacaaggt 440

<210> 179
 <211> 443
 <212> DNA
 <213> Homo sapien

<400> 179
 gaattcggca ccagcggggg gctacggcgg cggctacggc ggcgtcctga ccgcgtccga 60
 cgggctgctg gcgggcaacg agaagctaac catgcagaac ctcaacgacc gcttggcctc 120
 ctacctggac aagggtgcgcg ccctggaggc ggccaacggc gagctagagg tgaagatccg 180
 cgactggtac cagaagcagg ggcctgggcc ctcccgcgac tacagccact actacacgac 240
 catccaggac ctgcgggaca agattcttgg tgccaccatt gagaactcca ggattgtcct 300
 gcagatcgac aacgcccgtc tggctgcaga tgacttccga accaagtttg agacggaaca 360
 ggctctgcgc atgagcgtgg aggccgacat caacggcctg cgcagggtgc tggatgagct 420
 gacctgggcc aggaccgacc tgg 443

<210> 180
 <211> 403
 <212> DNA
 <213> Homo sapien

<400> 180
 gaattcggca cgaggttatg agagtcgact tcaatgttcc tatgaagaac aaccagataa 60
 caaacaacca gaggattaag gctgctgtcc caagcatcaa attctgcttg gacaatggag 120
 ccaagtcggg agtccttatg agccacctag gccggcctga tgggtgtgcc atgcttgaca 180
 agtactcctt agagccagtt gctgtagaac tcagatctct gctgggcaag gatgttctgt 240
 tcttgaagga ctgtgtaggc ccagaagtgg agaaagcctg tgccaacca gctgctgggt 300

Ser Gln Lys His Ser Leu Thr Ser Gln Ser Gln Ile Ser Pro Lys Ser
 260 265 270
 Trp Gly Val Ala Thr Ala Ser Leu Ile Pro Asn Asp Gln Leu Leu Pro
 275 280 285
 Arg Lys Leu Asn Thr Glu Pro Lys Asp Val Pro Lys Pro Val His Gln
 290 295 300
 Pro Val Gly Ser Ser Ser Thr Leu Pro Lys Asp Pro Val Leu Arg Lys
 305 310 315 320
 Glu Lys Leu Gln Asp Leu Met Thr Gln Ile Gln Gly Thr Cys Asn Phe
 325 330 335
 Met Gln Glu Ser Val Leu Asp Phe Asp Lys Pro Ser Ser Ala Ile Pro
 340 345 350
 Thr Ser Gln Pro Pro Ser Ala Thr Pro Gly Ser Pro Val Ala Ser Lys
 355 360 365
 Glu Gln Asn Leu Ser Ser Gln Ser Asp Phe Leu Gln Glu Pro Leu Gln
 370 375 380
 Val Phe Asn Val Asn Ala Pro Leu Pro Pro Arg Lys Glu Gln Glu Ile
 385 390 395 400
 Lys Glu Ser Pro Tyr Ser Pro Gly Tyr Asn Gln Ser Phe Thr Thr Ala
 405 410 415
 Ser Thr Gln Thr Pro Pro Gln Cys Gln Leu Pro Ser Ile His Val Glu
 420 425 430
 Gln Thr Val His Ser Gln Glu Thr Ala Ala Asn Tyr His Pro Asp Gly
 435 440 445
 Thr Ile Gln Val Ser Asn Gly Ser Leu Ala Phe Tyr Pro Ala Gln Thr
 450 455 460
 Asn Val Phe Pro Arg Pro Thr Gln Pro Phe Val Asn Ser Arg Gly Ser
 465 470 475 480
 Val Arg Gly Cys Thr Arg Gly Gly Arg Leu Ile Thr Asn Ser Tyr Arg
 485 490 495
 Ser Pro Gly Gly Tyr Lys Gly Phe Asp Thr Tyr Arg Gly Leu Pro Ser
 500 505 510
 Ile Ser Asn Gly Asn Tyr Ser Gln Leu Gln Phe Gln Ala Arg Glu Tyr
 515 520 525
 Ser Gly Ala Pro Tyr Ser Gln Arg Asp Asn Phe Gln Gln Cys Tyr Lys
 530 535 540
 Arg Gly Gly Thr Ser Gly Gly Pro Arg Ala Asn Ser Arg Ala Gly Trp
 545 550 555 560
 Ser Asp Ser Ser Gln Val Ser Ser Pro Glu Arg Asp Asn Glu Thr Phe
 565 570 575
 Asn Ser Gly Asp Ser Gly Gln Gly Asp Ser Arg Ser Met Thr Pro Val
 580 585 590
 Asp Val Pro Val Thr Asn Pro Ala Ala Thr Ile Leu Pro Val His Val
 595 600 605
 Tyr Pro Leu Pro Gln Gln Met Arg Val Ala Phe Ser Ala Ala Arg Thr
 610 615 620
 Ser Asn Leu Ala Pro Gly Thr Leu Asp Gln Pro Ile Val Phe Asp Leu
 625 630 635 640
 Leu Leu Asn Asn Leu Gly Glu Thr Phe Asp Leu Gln Leu Gly Arg Phe
 645 650 655
 Asn Cys Pro Val Asn Gly Thr Tyr Val Phe Ile Phe His Met Leu Lys
 660 665 670
 Leu Ala Val Asn Val Pro Leu Tyr Val Asn Leu Met Lys Asn Glu Glu
 675 680 685

Val Leu Val Ser Ala Tyr Ala Asn Asp Gly Ala Pro Asp His Glu Thr
 690 695 700
 Ala Ser Asn His Ala Ile Leu Gln Leu Phe Gln Gly Asp Gln Ile Trp
 705 710 715 720
 Leu Arg Leu His Arg Gly Ala Ile Tyr Gly Ser Ser Trp Lys Tyr Ser
 725 730 735
 Thr Phe Ser Gly Tyr Leu Leu Tyr Gln Asp
 740 745

<210> 186

<211> 705

<212> PRT

<213> Homo sapien

<400> 186

Ala Leu Leu Asn Val Arg Gln Pro Pro Ser Thr Thr Thr Phe Val Leu
 1 5 10 15
 Asn Gln Ile Asn His Leu Pro Pro Leu Gly Ser Thr Ile Val Met Thr
 20 25 30
 Lys Thr Pro Pro Val Thr Thr Asn Arg Gln Thr Ile Thr Leu Thr Lys
 35 40 45
 Phe Ile Gln Thr Thr Ala Ser Thr Arg Pro Ser Val Ser Ala Pro Thr
 50 55 60
 Val Arg Asn Ala Met Thr Ser Ala Pro Ser Lys Asp Gln Val Gln Leu
 65 70 75 80
 Lys Asp Leu Leu Lys Asn Asn Ser Leu Asn Glu Leu Met Lys Leu Lys
 85 90 95
 Pro Pro Ala Asn Ile Ala Gln Pro Val Ala Thr Ala Ala Thr Asp Val
 100 105 110
 Ser Asn Gly Thr Val Lys Lys Glu Ser Ser Asn Lys Glu Gly Ala Arg
 115 120 125
 Met Trp Ile Asn Asp Met Lys Met Arg Ser Phe Ser Pro Thr Met Lys
 130 135 140
 Val Pro Val Val Lys Glu Asp Asp Glu Pro Glu Glu Glu Asp Glu Glu
 145 150 155 160
 Glu Met Gly His Ala Glu Thr Tyr Ala Glu Tyr Met Pro Ile Lys Leu
 165 170 175
 Lys Ile Gly Leu Arg His Pro Asp Ala Val Val Glu Thr Ser Ser Leu
 180 185 190
 Ser Ser Val Thr Pro Pro Asp Val Trp Tyr Lys Thr Ser Ile Ser Glu
 195 200 205
 Glu Thr Ile Asp Asn Gly Trp Leu Ser Ala Leu Gln Leu Glu Ala Ile
 210 215 220
 Thr Tyr Ala Ala Gln Gln His Glu Thr Phe Leu Pro Asn Gly Asp Arg
 225 230 235 240
 Ala Gly Phe Leu Ile Gly Asp Gly Ala Gly Val Gly Lys Gly Arg Thr
 245 250 255
 Ile Ala Gly Ile Ile Tyr Glu Asn Tyr Leu Leu Ser Arg Lys Arg Ala
 260 265 270
 Leu Trp Phe Ser Val Ser Asn Asp Leu Lys Tyr Asp Ala Glu Arg Asp
 275 280 285
 Leu Arg Asp Ile Gly Ala Lys Asn Ile Leu Val His Ser Leu Asn Lys
 290 295 300
 Phe Lys Tyr Gly Lys Ile Ser Ser Lys His Asn Gly Ser Val Lys Lys

0596720-053000

$\langle 210 \rangle$	187
$\langle 211 \rangle$	595

<213> Homo sapien

<400> 187

Glu 1	Ser	Pro	Arg	His	Arg	Gly	Glu	Gly	Gly	Gly	Glu	Trp	Gly	Pro	Gly
Val	Pro	Arg	Glu	Arg	Arg	Glu	Ser	Ala	Gly	Glu	Trp	Gly	Ala	Asp	Thr
Pro	Lys	Glu	Gly	Gly	Glu	Ser	Ala	Gly	Glu	Trp	Gly	Ala	Glu	Val	Pro
Arg	Gly	Arg	Gly	Glu	Gly	Ala	Gly	Glu	Trp	Gly	Pro	Asp	Thr	Pro	Lys
Glu	Arg	Gly	Gln	Gly	Val	Arg	Glu	Trp	Gly	Pro	Glu	Ile	Pro	Gln	Glu
His	Gly	Glu	Ala	Thr	Arg	Asp	Trp	Ala	Leu	Glu	Ser	Pro	Arg	Ala	Leu
Gly	Glu	Asp	Ala	Arg	Glu	Leu	Gly	Ser	Ser	Pro	His	Asp	Arg	Gly	Ala
Ser	Pro	Arg	Asp	Leu	Ser	Gly	Glu	Ser	Pro	Cys	Thr	Gln	Arg	Ser	Gly
Leu	Leu	Pro	Glu	Arg	Arg	Gly	Asp	Ser	Pro	Trp	Pro	Pro	Trp	Pro	Ser
Pro	Gln	Glu	Arg	Asp	Ala	Gly	Thr	Arg	Asp	Arg	Glu	Glu	Ser	Pro	Arg
Asp	Trp	Gly	Gly	Ala	Glu	Ser	Pro	Arg	Gly	Trp	Glu	Ala	Gly	Pro	Arg
Glu	Trp	Gly	Pro	Ser	Pro	Ser	Gly	His	Gly	Asp	Gly	Pro	Arg	Arg	Arg
Pro	Arg	Lys	Arg	Arg	Gly	Arg	Lys	Gly	Arg	Met	Gly	Arg	Gln	His	Glu
Ala	Ala	Ala	Thr	Ala	Ala	Thr	Ala	Ala	Thr	Ala	Thr	Gly	Gly	Thr	Ala
Glu	Glu	Ala	Gly	Ala	Ser	Ala	Pro	Glu	Ser	Gln	Ala	Gly	Gly	Gly	Pro
Arg	Gly	Arg	Ala	Arg	Gly	Pro	Arg	Gln	Gln	Gly	Arg	Arg	Arg	His	Gly
Thr	Gln	Arg	Arg	Arg	Gly	Pro	Pro	Gln	Ala	Arg	Glu	Glu	Gly	Pro	Arg
Asp	Ala	Thr	Thr	Ile	Leu	Gly	Leu	Gly	Thr	Pro	Ser	Gly	Glu	Gln	Arg
Ala	Asp	Gln	Ser	Gln	Ala	Leu	Pro	Ala	Leu	Ala	Gly	Ala	Ala	Ala	Ala
His	Ala	His	Ala	Ile	Pro	Gly	Ala	Gly	Pro	Ala	Ala	Ala	Ala	Pro	Val
Gly	Arg	Gly	Arg	Arg	Gly	Gly	Trp	Arg	Gly	Gly	Arg	Arg	Gly	Gly	Ser
Ala	Gly	Ala	Gly	Gly	Gly	Gly	Arg	Gly	Gly	Arg	Gly	Arg	Gly	Arg	Gly
Gly	Gly	Arg	Gly	Gly	Gly	Gly	Ala	Gly	Arg	Gly	Gly	Gly	Gly	Ala	Gly
Pro	Arg	Glu	Gly	Ala	Ser	Ser	Pro	Gly	Ala	Arg	Arg	Gly	Glu	Gln	Arg
Arg	Arg	Gly	Arg	Gly	Pro	Pro	Ala	Ala	Gly	Ala	Ala	Gln	Val	Ser	Ala

165 170 175
 Ile Ser Gln Leu Glu Gln Lys Val Arg Glu Ser Glu Leu Gln Val His
 180 185 190
 Ser Ala Leu Leu Gly Arg Pro Ala Pro Phe Gly Asp Val Cys Leu Leu
 195 200 205
 Arg Leu Gln Glu Leu Gln Arg Glu Asn Thr Phe Leu Arg Ala Gln Phe
 210 215 220
 Ala Gln Lys Thr Glu Ala Leu Ser Lys Glu Lys Met Glu Leu Glu Lys
 225 230 235 240
 Lys Leu Ser Ala Ser Glu Val Glu Ile Gln Leu Ile Arg Glu Ser Leu
 245 250 255
 Lys Val Thr Leu Gln Lys His Ser Glu Glu Gly Lys Lys Gln Glu Glu
 260 265 270
 Arg Val Lys Gly Arg Asp Lys His Ile Asn Asn Leu Lys Lys Lys Cys
 275 280 285
 Gln Lys Glu Ser Glu Gln Asn Arg Glu Lys Gln Gln Arg Ile Glu Thr
 290 295 300
 Leu Glu Arg Tyr Leu Ala Asp Leu Pro Thr Leu Glu Asp His Gln Lys
 305 310 315 320
 Gln Thr Glu Gln Leu Lys Asp Ala Glu Leu Lys Asn Thr Glu Leu Gln
 325 330 335
 Glu Arg Val Ala Glu Leu Glu Thr Leu Leu Glu Asp Thr Gln Ala Thr
 340 345 350
 Cys Arg Glu Lys Glu Val Gln Leu Glu Ser Leu Arg Gln Arg Glu Ala
 355 360 365
 Asp Leu Ser Ser Ala Arg His Arg
 370 375

<210> 189

<211> 160

<212> PRT

<213> Homo sapien

<400> 189

Met Leu Glu Ala His Arg Arg Gln Arg His Pro Phe Leu Leu Leu Gly
 1 5 10 15
 Thr Thr Ala Asn Arg Thr Gln Ser Leu Asn Tyr Gly Cys Ile Val Glu
 20 25 30
 Asn Pro Gln Thr His Glu Val Leu His Tyr Val Glu Lys Pro Ser Thr
 35 40 45
 Phe Ile Ser Asp Ile Ile Asn Cys Gly Ile Tyr Leu Phe Ser Pro Glu
 50 55 60
 Ala Leu Lys Pro Leu Arg Asp Val Phe Gln Arg Asn Gln Gln Asp Gly
 65 70 75 80
 Gln Leu Glu Asp Ser Pro Gly Leu Trp Pro Gly Ala Gly Thr Ile Arg
 85 90 95
 Leu Glu Gln Asp Val Phe Ser Ala Leu Ala Gly Gln Gly Gln Ile Tyr
 100 105 110
 Val His Leu Thr Asp Gly Ile Trp Ser Gln Ile Lys Ser Ala Gly Ser
 115 120 125
 Ala Leu Tyr Ala Ser Arg Leu Tyr Leu Ser Arg Tyr Gln Asp Thr His
 130 135 140
 Pro Glu Arg Leu Ala Lys His Thr Pro Gly Gly Pro Trp Ile Arg Gly
 145 150 155 160

000260"026960

Ser Phe Ile Gly Gln Asn Lys Ser Phe Phe Gly Pro Leu Glu Leu Val
 165 170 175
 Glu Lys Leu Cys Pro Glu Ala Ser Asp Ile Ala Thr Ser Val Arg Asn
 180 185 190
 Leu Pro Glu Leu Lys Thr Ala Val Gly Arg Gly Arg Ala Trp Leu Tyr
 195 200 205
 Leu Ala Leu Met Gln Lys Lys Leu Ala Asp Tyr Leu Lys Val Leu Ile
 210 215 220
 Asp Asn Lys His Leu Leu Ser Glu Phe Tyr Glu Pro Glu Ala Leu Met
 225 230 235 240
 Met Glu Glu Glu Gly Met Val Ile Val Gly Leu Leu Val Gly Leu Asn
 245 250 255
 Val Leu Asp Ala Asn Leu Cys Leu Lys Gly Glu Asp Leu Asp Ser Gln
 260 265 270
 Val Gly Val Ile Asp Phe Ser Leu Tyr Leu Lys Asp Val Gln Asp Leu
 275 280 285
 Asp Gly Gly Lys Glu His Glu Arg Ile Thr Asp Val Leu Asp Gln Lys
 290 295 300
 Asn Tyr Val Glu Glu Leu Asn Arg His Leu Ser Cys Thr Val Gly Asp
 305 310 315 320
 Leu Gln Thr Lys Ile Asp Gly Leu Glu Lys Thr Asn Ser Lys Leu Gln
 325 330 335
 Glu Glu Leu Ser Ala Ala Thr Asp Arg Ile Cys Ser Leu Gln Glu Glu
 340 345 350
 Gln Gln Gln Leu Arg Glu Gln Asn Glu Leu Ile Arg Glu Arg Ser Glu
 355 360 365
 Lys Ser Val Glu Ile Thr Lys Gln Asp Thr Lys Val Glu Leu Glu Thr
 370 375 380
 Tyr Lys Gln Thr Arg Gln Gly Leu Asp Glu Met Tyr Ser Asp Val Trp
 385 390 395 400
 Lys Gln Leu Lys Glu Glu Lys Lys Val Arg Leu Glu Leu Glu Lys Glu
 405 410 415
 Leu Glu Leu Gln Ile Gly Met Lys Thr Glu Met Glu Ile Ala Met Lys
 420 425 430
 Leu Leu Glu Lys Asp Thr His Glu Lys Gln Asp Thr Leu Val Ala Leu
 435 440 445
 Arg Gln Gln Leu Glu Glu Val Lys Ala Ile Asn Leu Gln Met Phe His
 450 455 460
 Lys Ala Gln Asn Ala Glu Ser Ser Leu Gln Gln Lys Asn Glu Ala Ile
 465 470 475 480
 Thr Ser Phe Glu Gly Lys Thr Asn Gln Val Met Ser Ser Met Lys Gln
 485 490 495
 Met Glu Glu Arg Leu Gln His Ser Glu Arg Ala Arg Gln Gly Ala Glu
 500 505 510
 Glu Arg Ser His Lys Leu Gln Gln Glu Leu Gly Gly Arg Ile Gly Ala
 515 520 525
 Leu Gln Leu Gln Leu Ser Gln Leu His Glu Gln Cys Ser Ser Leu Glu
 530 535 540
 Lys Glu Leu Lys Ser Glu Lys Glu Gln Arg Gln Ala Leu Gln Arg Glu
 545 550 555 560
 Leu Gln His Glu Lys Asp Thr Ser Ser Leu Leu Arg Met Glu Leu Gln
 565 570 575
 Gln Val Glu Gly Leu Lys Lys Glu Leu Arg Glu Leu Gln Asp Glu Lys
 580 585 590

Ala Thr Gln Glu Glu Leu Lys Lys Ala Tyr Arg Lys Leu Ala Leu Lys
 20 25 30
 Tyr His Pro Asp Lys Asn Pro Asn Glu Gly Glu Lys Phe Lys Gln Ile
 35 40 45
 Ser Gln Ala Tyr Glu Val Leu Ser Asp Ala Lys Lys Arg Glu Leu Tyr
 50 55 60
 Asp Lys Gly Gly Glu Gln Ala Ile Lys Glu Gly Gly Ala Gly Gly Gly
 65 70 75 80
 Phe Gly Ser Pro Met Asp Ile Phe Asp Met Phe Phe Gly Gly Gly Gly
 85 90 95
 Arg Met Gln Arg Glu Arg Arg Gly Lys Asn Val Val His Gln Leu Ser
 100 105 110
 Val Thr Leu Glu Asp Leu Tyr Asn Gly Ala Thr Arg Lys Leu Ala
 115 120 125

<210> 200

<211> 90

<212> PRT

<213> Homo sapien

<400> 200

Met Ala Cys Pro Leu Asp Gln Ala Ile Gly Leu Leu Val Ala Ile Phe
 1 5 10 15
 His Lys Tyr Ser Gly Arg Glu Gly Asp Lys His Thr Leu Ser Lys Lys
 20 25 30
 Glu Leu Lys Glu Leu Ile Gln Lys Glu Leu Thr Ile Gly Ser Lys Leu
 35 40 45
 Gln Asp Ala Glu Ile Ala Arg Leu Met Glu Asp Leu Asp Arg Asn Lys
 50 55 60
 Asp Gln Glu Val Asn Phe Gln Glu Tyr Val Thr Phe Leu Gly Ala Leu
 65 70 75 80
 Ala Leu Ile Tyr Asn Glu Ala Leu Lys Gly
 85 90

<210> 201

<211> 120

<212> PRT

<213> Homo sapien

<400> 201

Met Glu Thr Pro Ser Gln Arg Arg Ala Thr Arg Ser Gly Ala Gln Ala
 1 5 10 15
 Ser Ser Thr Pro Leu Ser Pro Thr Arg Ile Thr Arg Leu Gln Glu Lys
 20 25 30
 Glu Asp Leu Gln Glu Leu Asn Asp Arg Leu Ala Val Tyr Ile Asp Arg
 35 40 45
 Val Arg Ser Leu Glu Thr Glu Asn Ala Gly Leu Arg Leu Arg Ile Thr
 50 55 60
 Glu Ser Glu Glu Val Val Ser Arg Glu Val Ser Gly Ile Lys Ala Ala
 65 70 75 80
 Tyr Glu Ala Glu Leu Gly Asp Ala Arg Lys Thr Leu Asp Ser Val Ala
 85 90 95
 Lys Glu Arg Ala Arg Leu Gln Leu Glu Leu Ser Lys Val Arg Glu Glu
 100 105 110

0002260"022950

<210> 207
 <211> 175
 <212> PRT
 <213> Homo sapien

<400> 207

Ile	Ile	Arg	Gln	Gln	Gly	Leu	Ala	Ser	Tyr	Asp	Tyr	Val	Arg	Arg	Arg
1				5					10				15		
Leu	Thr	Ala	Glu	Asp	Leu	Phe	Glu	Ala	Arg	Ile	Ile	Ser	Leu	Glu	Thr
			20					25					30		
Tyr	Asn	Leu	Leu	Arg	Glu	Gly	Thr	Arg	Ser	Leu	Arg	Glu	Ala	Leu	Glu
	35						40					45			
Ala	Glu	Ser	Ala	Trp	Cys	Tyr	Leu	Tyr	Gly	Thr	Gly	Ser	Val	Ala	Gly
	50					55					60				
Val	Tyr	Leu	Pro	Gly	Ser	Arg	Gln	Thr	Leu	Ser	Ile	Tyr	Gln	Ala	Leu
65					70					75				80	
Lys	Lys	Gly	Leu	Leu	Ser	Ala	Glu	Val	Ala	Arg	Leu	Leu	Leu	Glu	Ala
			85						90					95	
Gln	Ala	Ala	Thr	Gly	Phe	Leu	Leu	Asp	Pro	Val	Lys	Gly	Glu	Arg	Leu
			100					105					110		
Thr	Val	Asp	Glu	Ala	Val	Arg	Lys	Gly	Leu	Val	Gly	Pro	Glu	Leu	His
		115					120					125			
Asp	Arg	Leu	Leu	Ser	Ala	Glu	Arg	Ala	Val	Thr	Gly	Tyr	Arg	Asp	Pro
	130					135					140				
Tyr	Thr	Glu	Gln	Thr	Ile	Ser	Leu	Phe	Gln	Ala	Met	Lys	Lys	Glu	Leu
145					150					155					160
Ile	Pro	Thr	Glu	Glu	Ala	Leu	Arg	Leu	Trp	Met	Pro	Ser	Trp	Pro	
			165						170					175	

<210> 208
 <211> 177
 <212> PRT
 <213> Homo sapien

<400> 208

Met	Ala	Ala	Gly	Val	Glu	Ala	Ala	Ala	Glu	Val	Ala	Ala	Thr	Glu	Ile
1				5					10				15		
Lys	Met	Glu	Glu	Glu	Ser	Gly	Ala	Pro	Gly	Val	Pro	Ser	Gly	Asn	Gly
		20						25					30		
Ala	Pro	Gly	Pro	Lys	Gly	Glu	Gly	Glu	Arg	Pro	Ala	Gln	Asn	Glu	Lys
	35						40					45			
Arg	Lys	Glu	Lys	Asn	Ile	Lys	Arg	Gly	Gly	Asn	Arg	Phe	Glu	Pro	Tyr
	50					55				60					
Ala	Asn	Pro	Thr	Lys	Arg	Tyr	Arg	Ala	Phe	Ile	Thr	Asn	Ile	Pro	Phe
65					70					75				80	
Asp	Val	Lys	Trp	Gln	Ser	Leu	Lys	Asp	Leu	Val	Lys	Glu	Lys	Val	Gly
			85						90					95	
Glu	Val	Thr	Tyr	Val	Glu	Leu	Leu	Met	Asp	Ala	Glu	Gly	Lys	Ser	Arg
		100						105					110		
Gly	Cys	Ala	Val	Val	Glu	Phe	Lys	Met	Glu	Glu	Ser	Met	Lys	Lys	Ala
		115					120					125			
Ala	Glu	Val	Leu	Asn	Lys	His	Ser	Leu	Ser	Gly	Arg	Pro	Leu	Lys	Val
	130					135					140				
Lys	Glu	Asp	Pro	Asp	Gly	Glu	His	Ala	Arg	Arg	Ala	Met	Gln	Lys	Val

00667170-092000

```
<210> 209
<211> 196
<212> PRT
<213> Homo sapien
```

```
<210> 210
<211> 156
<212> PRT
<213> Homo sapien
```

			<400>	210											
Lys	Leu	Thr	Ile	Glu	Ser	Thr	Pro	Phe	Asn	Val	Ala	Glu	Gly	Lys	Glu
1				5					10					15	
Val	Leu	Leu	Leu	Ala	His	Asn	Leu	Pro	Gln	Asn	Arg	Ile	Gly	Tyr	Ser
			20					25					30		
Trp	Tyr	Lys	Gly	Glu	Arg	Val	Asp	Gly	Asn	Ser	Leu	Ile	Val	Gly	Tyr
		35					40					45			
Val	Ile	Gly	Thr	Gln	Gln	Ala	Thr	Pro	Gly	Pro	Ala	Tyr	Ser	Gly	Arg
	50					55					60				
Glu	Thr	Ile	Tyr	Pro	Asn	Ala	Ser	Leu	Leu	Ile	Gln	Asn	Val	Thr	Gln

<212> PRT

<213> Homo sapien

<400> 215

```

Met Ala Thr Leu Lys Glu Lys Leu Ile Ala Pro Val Ala Glu Glu Glu
 1      5      10      15
Ala Thr Val Pro Asn Asn Lys Ile Thr Val Val Gly Val Gly Gln Val
 20      25      30
Gly Met Ala Cys Ala Ile Ser Ile Leu Gly Lys Ser Leu Ala Asp Glu
 35      40      45
Leu Ala Leu Val Asp Val Leu Glu Asp Lys Leu Lys Gly Glu Met Met
 50      55      60
Asp Leu Gln His Gly Ser Leu Phe Leu Gln Thr Pro Lys Ile Val Ala
 65      70      75      80
Asp Lys Asp Tyr Ser Val Thr Ala Asn Ser Lys Ile Val Val Val Thr
 85      90      95
Ala Gly Val Arg Gln Gln Glu Gly Glu Ser Arg Leu Asn Leu Val Gln
 100     105     110
Arg Asn Val Asn Val Phe Lys Phe Ile Ile Pro Gln Ile Val Lys Tyr
 115     120     125
Ser Pro Asp Cys Ile Ile Ile Val Val Ser Asn Pro Val Asp Ile Leu
 130     135     140
Thr Tyr Val Thr
145

```

<210> 216

<211> 527

<212> PRT

<213> Homo sapien

<400> 216

```

Gln Arg Ala Pro Gly Ile Glu Glu Lys Ala Ala Glu Asn Gly Ala Leu
 1      5      10      15
Gly Ser Pro Glu Arg Glu Glu Lys Val Leu Glu Asn Gly Glu Leu Thr
 20      25      30
Pro Pro Arg Arg Glu Glu Lys Ala Leu Glu Asn Gly Glu Leu Arg Ser
 35      40      45
Pro Glu Ala Gly Glu Lys Val Leu Val Asn Gly Gly Leu Thr Pro Pro
 50      55      60
Lys Ser Glu Asp Lys Val Ser Glu Asn Gly Gly Leu Arg Phe Pro Arg
 65      70      75      80
Asn Thr Glu Arg Pro Pro Glu Thr Gly Pro Trp Arg Ala Pro Gly Pro
 85      90      95
Trp Glu Lys Thr Pro Glu Ser Trp Gly Pro Ala Pro Thr Ile Gly Glu
 100     105     110
Pro Ala Pro Glu Thr Ser Leu Glu Arg Ala Pro Ala Pro Ser Ala Val
 115     120     125
Val Ser Ser Arg Asn Gly Gly Glu Thr Ala Pro Gly Pro Leu Gly Pro
 130     135     140
Ala Pro Lys Asn Gly Thr Leu Glu Pro Gly Thr Glu Arg Arg Ala Pro
 145     150     155     160
Glu Thr Gly Gly Ala Pro Arg Ala Pro Gly Ala Gly Arg Leu Asp Leu
 165     170     175
Gly Ser Gly Gly Arg Ala Pro Val Gly Thr Gly Thr Ala Pro Gly Gly

```

000000"000000

ggaagcagga	cagccaactt	cagaaggcca	gggaggacat	ctttatggag	accctgaaa	300
acatcgtgga	gtattacaac	gacagtaacg	ggtctcacgt	attgcaggga	aggtttggtt	360
gtgagatcga	gaataacaga	agcagcggag	cattctggaa	atattactat	gatggaaagg	420
actacattga	attcaacaaa	gaaatcccag	cctgggtccc	cttcga		466

<210> 218

<211> 381

<212> DNA

<213> Homo sapien

<400> 218

gagtttcctt	cgcaagttca	tgtggggtac	cttcccaggc	tgcctggctg	accagctggt	60
tttaaagcgc	cggggttaacc	agttggagat	ctgtgccgtg	gtcctgaggc	agttgtctcc	120
acacaagtac	tacttcctcg	tgggctacag	tgaaactttg	ctgtcctact	tttacaaatg	180
tctgtgcgga	ctccacctcc	aaactgtgcc	ctcaaagggt	gtgtataagt	acctctagaa	240
caatccccctt	ttttccatca	agctgtagcc	tgcagagaat	ggaaacgtgg	gaaagggaatg	300
gtatgtgggg	gaaatgcata	ccctcagagg	actgaggcat	agtctctcat	ctgctattga	360
ataaagacct	tctatcttgt	a				381

<210> 219

<211> 1293

<212> DNA

<213> Homo sapien

<400> 219

gaggggaggc	gcatggcggg	gatggcgctg	gcgcgggcct	ggaagcagat	gtcctggttc	60
tactaccagt	acctgctggt	cacggcgctc	tacatgctgg	agccctggga	gcggaagggtg	120
ttcaattcca	tgctggtttc	cattgtgggg	atggcactat	acacaggata	cgtcttcatg	180
ccccagcaca	tcatggcgat	attgcactac	tttgaaatcg	tacaatgacc	aagatgcgac	240
caggatcaga	ggttccttgg	ggaagaccca	ccctacgaag	ttggaatgag	accatcagat	300
gtgataagaa	actcttctag	atgtcaacat	aaccaacctt	ataaagacta	aaattcatga	360
gtagaacagg	aaaatcatcc	tgactcatgt	gttgtgttct	ttatTTTTaa	ttttcaaaga	420
ggctcttgta	tagcagtttt	tgtctatttt	aacattgtag	tcatttgtac	tttgatatca	480
gtattttctt	aacctttgtg	actgtttcaa	tattaccccc	gtgaaagctt	ttcttaatgt	540
aactttgagt	acattttaat	tgccttctat	ttttaaaact	caaaatcatt	agttgggctt	600
tactgttctt	gctattgtat	ggcatatata	tctgcctgga	tatatttcta	ctcttgacca	660
aagttttgta	aagaacaata	taagattttcg	ggtaggggta	tggggaggga	agatatttta	720
ttgagaacta	cttaacaaaa	gatttatctg	taagcttgaa	ctcaggagta	cagtttttagc	780
tatctagact	ctaacagctt	ttgcttttaa	attattaaag	tgtttcttaa	tgaaaaagaa	840
aagatcttgc	taaagttaaa	ataaggaaca	tttcaccttt	taaaatattta	attcttatgt	900
ggacttattt	ccagaaaact	ttggtgataa	ttcttgagac	aaaagggtgt	taagtagcat	960
tattatgtaa	tgcttatata	ccatagagtt	tttaatagaa	gagaaatcca	tttcctccga	1020
gggtcactat	taacaatgta	cttccttaaa	tttagtttaa	tgattgtaat	gggtgctgca	1080
tttgcacatt	gcattaaagt	atgatgagac	gaattgttgt	taaaaattat	agcaaaaaga	1140
aatgtaaact	tggttaaaat	cctttcactc	tttgtattgt	tttttttaag	gtttttattc	1200
cttaaatgta	aaatgactac	ctaatttttt	gatgtaaata	cattaaattc	aaagagaaaa	1260
aaaatcaaaa	aaaaaaaaaa	aaaaaaactc	gag			1293

<210> 220

<211> 983

<212> DNA

<213> Homo sapien

<400> 220

<400> 223

gaggcaagg	atattgcttta	gtgcctatta	tagttaattc	ttcaactcca	aagtctaaaa	60
cagttgaatc	tgctgaagga	aaatctgaag	aagtaaatga	aacattagtt	ataccactg	120
aggaagcaga	aatggaagaa	agtggacgaa	gtgcaactcc	tgtaactgt	gaacagcctg	180
atatcttgg	ttcttctaca	ccaataaatg	aaggacagac	tgtgttagac	aaggtggctg	240
agcagtgtga	acctgctgaa	agtcagccag	aagcacttct	gagaggaaga	tgtttgcaag	300
gtaactctaa	cagttg					316

<210> 224

<211> 1583

<212> DNA

<213> Homo sapien

<400> 224

cagaccacgt	ctgcctctgc	cgctctagcc	ctgcgcccc	gcccggccgc	ggcacctcgc	60
cctcgccgcc	gctaggtcgg	ccggctccgc	ccggctgccc	cctaggatga	atatcatgga	120
cttcaacgtg	aagaagctgg	cgcccgacgc	aggcaccttc	ctcagtcgcg	ccgtgcagtt	180
cacagaagaa	aagcttgccc	aggctgagaa	gacagaattg	gatgctcact	tagagaacct	240
ccttagcaaa	gctgaatgta	ccaaaatatg	gacagaaaaa	ataatgaaac	aaactgaagt	300
gttattgcag	ccaaatccaa	atgccaggat	agaagaattt	gtttatgaga	aactggatag	360
aaaagctcca	agtcgtataa	acaaccacga	acttttgcca	caatatatga	ttgatgcagg	420
gactgagttt	ggcccaggaa	cagcttatgg	taatgccttt	attaaatgtg	gagaaaccca	480
aaaaagaatt	ggaacagcag	acagagaact	gattcaaacg	tcagccttaa	attttcttac	540
tcctttaaga	aactttatag	aaggagatta	caaaacaatt	gctaaagaaa	ggaaactatt	600
gcaaaataag	agactggatt	tggtatgctgc	aaaaacgaga	ctaaaaaagg	caaaagctgc	660
agaaactaga	aattcatctg	aacaggaatt	aagaataact	caaagtgaat	ttgatcgtea	720
agcagagatt	accagacttc	tgctagaggg	aatcagcagt	acacatgccc	atcaccttcg	780
ctgtctgaat	gactttgtag	aagcccagat	gacttactat	gcacagtgtt	accagtatat	840
gttggacctc	cagaaacaac	tggaagttt	tccatccaat	tatcttagta	acaacaatca	900
gacttctgtg	acacctgtac	catcagtttt	accaaattgc	attggttctt	ctgccatggc	960
ttcaacaagt	ggcctagtaa	tcacctctcc	ttccaacctc	agtgcactta	aggagtgtag	1020
tggcagcaga	aaggccaggg	ttctctatga	ttatgatgca	gcaaacagta	ctgaattatc	1080
acttctggca	gatgaggtga	tcactgtgtt	cagtgttgtt	ggaatggatt	cagactggct	1140
aatgggggaa	aggggaaacc	agaagggcaa	ggtgccaat	acctacttag	aactgctcaa	1200
ttaagttagt	ggactatgga	aagggtgccc	atcatgactt	tgtattttata	tacaattaac	1260
tctaaataaa	gcagggttaag	tatcttccat	gttaatgtgt	taagagactg	aaaataaccag	1320
ccatcagaaa	ctggcctttt	tgccaataaa	gttgcattgt	aaatatttca	ttacagaatt	1380
tatgttagag	ctttcatgcc	aagaatgttt	tcttacaaaa	ttctcttttt	attgaggttt	1440
cactaataag	cagcttctac	ttttgagcct	caacttaaa	cagaactgtt	ttttactgga	1500
tttttcatta	acagcaagct	ttttttttta	tgtaaaataa	atctatttgt	aattgaaaaa	1560
aaaaaaaaaa	aaaaaaactc	gag				1583

<210> 225

<211> 491

<212> DNA

<213> Homo sapien

<400> 225

gaacaacatc	atcttgaatc	actagataga	ctcttgacgg	aaagcaaagg	ggaaatgaaa	60
aaggaaaata	tgaagaaaga	tgaagcttta	aaagcattac	agaaccaagt	atctgaagaa	120
acaatcaagg	ttaggcaact	agattcagca	ttggaaattt	gtaaggaaga	acttgtcttg	180
catttgaatc	aattggaagg	aaataaggaa	aagtttgaaa	aacagttaaa	gaagaaatct	240
gaagaggtat	attgtttaca	gaaagagcta	aagataaaaa	atcacagtct	tcaagagact	300
tctgagcaaa	acgttattct	acagcatact	cttcagcaac	agcagcaaat	gttacaacaa	360

gagacaatta gaaatggaga gctagaagat actcaaacta aacttgaaaa acaggtgtca 420
 aaactggaac aagaacttca aaaacaaagg gaaagttcag ctgaaaagtt gagaaaaatg 480
 gaggagaaat g 491

<210> 226

<211> 483

<212> DNA

<213> Homo sapien

<400> 226

cagccgcacg ccgcggagca ggggctcgga ggtcccggga ttacggtgct cgagcacgct 60
 ggtgggaaag gacccgggac ttgaacagtg ttgtgcgcg ccatgcaggt ctccagcctc 120
 aatgaggtga agattttacag cctcagctgc ggcaagtccc ttcctgagtg gctttctgat 180
 aggaagaaga gagcgctaca gaagaaagat gtagatgtcc gtaggagaaat tgaacttatt 240
 caggactttg aaatgcctac tgttgtgtacc actattaagg tgtcaaaaga tggacagtac 300
 atttttagcaa ctggaacata taaacctcgg gttcgatgtt atgacaccta tcaattatcc 360
 ttgaagtttg aaaggtgttt agattcagaa gttgtcacct ttgaaatfff gtctgatgac 420
 tactcaaaga ttgtcttctt acataatgat agatacattg aatttcattc gcaatcaggt 480
 ttt 483

<210> 227

<211> 486

<212> DNA

<213> Homo sapien

<400> 227

gagcctcgct aagctccgac tctgggcggc accgggcgtc ccacgatgcc gaagaacaag 60
 aagcggaaaca ctccccaccg cggtagcagt gctggcgcg gcgggtcagg agcagccgca 120
 gcgacggcgg cgacagcagg tggccagcat cgaaatgttc agccttttag tgatgaagat 180
 gcatcaattg aaacagtgag ccattgcagt gggtatagcg atccttccag ttttgctgaa 240
 gatggaccag aagtccttga tgaggaagga actcaagaag acctagagta caagttgaag 300
 ggattaattg acctaacctt ggataagagt gcgaagacaa ggcaagcagc tcttgaaggt 360
 attaaaaatg cactggcttc aaaaatgctg tatgaattta ttctggaaag gagaatgact 420
 ttaactgata gcattgaacg ctgcctgaaa aaaggtgaaga gtgatgagca acgtgcagct 480
 gcagcg 486

<210> 228

<211> 494

<212> DNA

<213> Homo sapien

<400> 228

gaggccagga ctccgggaat gcgagcaggc cccttattct ccagtgggcc tcggtctgtc 60
 cccacagcgg ccggtcagg gttgcccgag ccccaaggcg gggggcggca ccgggggtgt 120
 gaaagggaca gaatgctttg acctccaagc tgttttaaat ctagtagata agccagatcc 180
 tgtgttgcca taagcccttg gccacattt aagtgggaat gcagctagct tggatgtctg 240
 aaactttgta agcgccttct gtctgaatcc tgaacacagg caccaagact actgaagaag 300
 ctcgctcattc ttgtgcaggg atagccacac aagcaaacat gtttgcaaaa cttgaaagaa 360
 agaaaattgc agaaagaaga cttgctgttc ttaagaggcc caggaagggtg ctacttagga 420
 atcccaccgg cttgtgaagc aagggaatca agtttgcctt caatggggaa cttgacttca 480
 ggaaaaatgaa cttt 494

<210> 229

<211> 465

<211> 278
 <212> DNA
 <213> Homo sapien

<400> 236
 gagattagca cctgtgaaca atgcgttctc tgatgacact ctgagcatgg accaacgcct 60
 tcttaagcta attctgcaaa atcacatatt gaaagtaaaa gttggcctta gcgacctcta 120
 caatggacag atactggaaa ccattggagg caaacaactc cgagtctttg tgtatcggac 180
 ggctatctgc atagaaaact catgcatggg gagaggaagc aagcagggaa ggaacggtgc 240
 cattcacata ttccgagaga tcatccaacc agcagaat 278

<210> 237
 <211> 322
 <212> DNA
 <213> Homo sapien

<400> 237
 cagggccgtg gcggaggagg agcgtgcac ggtggagcgt cgggccgacc tcacctacgc 60
 ggagttcgtg cagcagtagc tgcgcccctg atcggggagg tcgcgtcctg ttcaccggcc 120
 cgtctgcccc gaccgcccac ggccgccttc ccttgacctc gcgcgcacgc gtggggctgg 180
 ggcgggcagg ctggcggtcc ggccctggccg cgactctgcc cttctttcca gaggttccgg 240
 gccctgtgct cccgcgacag gttgctgggt tcggttgggg acagagtggg ccggtgagca 300
 ccgccaacac ctactcctac ct 322

<210> 238
 <211> 613
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (399)
 <223> n=A,T,C or G

<400> 238
 gaattcggca ccagccttct tggatcagga ccagtctcca ccccgtttct acagtggaga 60
 tcagcctcct tcttatcttg gtgcaagtgt ggataaactc catcaccctt tagaatttgc 120
 agacaaatct cccacacctc ctaatttacc tagcgataaa atctaccctc cttctgggtc 180
 ccccgaaagag aataccagca cagccaccat gacttacatg acaactactc cagcaacagc 240
 ccaaattgagc accaaggaag ccagctggga tgtggctgaa caaccacca ctgctgattt 300
 tgctgctgcc aacttcagc gcacgcacag aactaatcgt ccccttcccc ctccgccttc 360
 ccagagatct gcagagcagc caccagttgt ggggcaggna caagcagcaa ccaatatagg 420
 attaaataat tcccacaagg ttcaaggagt agttccagtt ccagagaggc cacctgaacc 480
 tcgagccatg gatgacctg cgtctgcctt catcagtgac agtgggtgctg ctgctgctca 540
 gtgtcccatg gctacagctg tccagccagg cctgcctgag aaagtgcggg acggtgcccc 600
 ggtcccgctg ctg 613

<210> 239
 <211> 613
 <212> DNA
 <213> Homo sapiens

<400> 239


```
<210> 243
<211> 591
<212> DNA
<213> Homo sapiens
```

```
<210> 244
<211> 594
<212> DNA
<213> Homo sapiens
```

```
<210> 245
<211> 615
<212> DNA
<213> Homo sapiens
```

<400> 245


```
<210> 249
<211> 416
<212> DNA
<213> Homo sapiens
```

```
<210> 250
<211> 504
<212> DNA
<213> Homo sapiens
```

```
<210> 251
<211> 607
<212> DNA
<213> Homo sapiens
```

<400>	251						
gatgaaaata	cacaatttta	ctagcaaatg	cctctactgt	aatcgctatt	tacccacaga	60	
tactctgctc	aaccatatgt	taattcatgg	tctgtcttgt	ccatattgcc	gttcaacttt	120	
caatgatgtg	gaaaagatgg	cgcacacat	gcggatggtt	cacattgatg	aagagatggg	180	
acctaaaaca	gattctactt	tgagttttga	tttgacattg	cagcagggta	gtcacactaa	240	
catccatctc	ctggtaacta	catacaatct	gagggatgcc	ccagctgaat	ctggtgctta	300	
ccatgcccaa	aataatcctc	cagttcctcc	aaagccacag	ccaagggttc	aggaaaaggc	360	
agatatccct	gtaaaaagtt	cacctcaagc	tgcagtgcc	tataaaaaag	atgttgggaa	420	
aaccctttgt	cctcttttgt	tttcaatcc	aaaaggacc	atatctgatg	cacttgca	480	
tcacttacga	gagaggcacc	aagttattca	gacggttcat	ccagttgaga	aaaagctcac	540	
tcacaaatgt	atccattgcc	ttggtgtgta	taccagcaac	atgaccgcct	caactatcac	600	
tctgcat						607	

<210> 252
 <211> 618
 <212> DNA
 <213> Homo sapiens

<400> 252
 gaattcgcac caggggtcct gctgggtcttc gcctttcttc tccgcttcta ccccgtcggc 60
 cgctgccact ggggtccctg gcccaccga catggcgcg gtgttgagca agtcctggag 120
 cgcacggagc tgaacaagct gcccaagtct gtccagaaca aacttgaaaa gttccttgct 180
 gatcagcaat ccgagatcga tggcctgaag gggcgcatg agaaatttaa ggtggagagc 240
 gaacaacagt attttgaaat agaaaagagg ttgtcccaca gtcaggagag acttgtgaat 300
 gaaacccgag agtgtcaaag cttgcggctt gagctagaga aactcaaca tcaactgaag 360
 gactaactg agaaaaacaa agaacttgaa attgctcagg atcgcaatat tgccattcag 420
 agccaattta caagaacaaa ggaagaatta gaagctgaga aaagagactt aattagaacc 480
 aatgagagac tatctcaaga acttgaatac ttaacagagg atgttaaacg tctgaatgaa 540
 aaacttaaag aaagcaatac acaaaagggt gaacttcagt taaaattgga tgaacttcaa 600
 gcttctgatg tttctggt 618

<210> 253
 <211> 1201
 <212> DNA
 <213> Homo sapiens

<400> 253
 gaattcggca ccagggtggc gagcgcggt gctgtgctgg ggcgagcagc ggggaccgtg 60
 tgtgagtttg gcatgatttg gtcccctggg attctgcctt agcaagaaa aagttggaaa 120
 tacttccttg aagaaaacta aaacaatata aaagccacag cttattgatt gcatgtcagc 180
 ccccttaca atattggacac atttcctagc ctatttccac ctggaggaga tagtaggctg 240
 aatcctgagc ctgagttcca aaatatgtta attgatgaaa gggtagcgtg tgaacatcat 300
 aacataatt atcaggctct gaaaattgaa caaaaagggt tgcaggaaga atatgtaaaa 360
 tcacaaaatg aacttaaacg tgtattaatt gaaaagcaag caagccagga aaaattccaa 420
 ctgctccttg aagacttaag gggagaatta gtagagaaag ctagagacat agaaaaaatg 480
 aaactgcagg tactaacacc acaaaaattg gaattggtta aagcccaact acaacaagaa 540
 ttagaagctc caatgcgaga acgttttcgg actcttgatg aagaagtgga aaggtagaga 600
 gctgagtata acaagctgcg ctacgagtat acatttctca agtcagagtt tgaacaccag 660
 aaagaagagt ttactcgggt ttcagaagaa gagaaaatga aatacaagtc agaggttgca 720
 cgactggaga aggacaaaga ggagctacat aaccagctgc ttagtggtga tcccacgaga 780
 gacagcaaac gaatggagca acttgttcga gaaaaaacc atttgcttca gaaattgaaa 840
 agtttagagg ctgaagtagc agaattaagg gctgagaaag aaaattctgg tgctcaggta 900
 gaaaatgtcc aaagaatata ggtgaggcag ttggctgaga tgcaggctac actcagatcc 960
 ttggaggctg aaaagcagtc agctaaacta caagctgagc gtttagaaaa agaactacaa 1020
 tcaagcaatg aacagaatac ctgcttaatc agcaaactgc atagagctga ccgagaaatc 1080
 agcacactgg ccagtgaagt gaaagagctt aaacatgcaa acaaactaga aataactgac 1140
 atcaaactgg aggcagcaag agctaagagt gagctcgaaa gagaaaggaa taagatccaa 1200
 a 1201

<210> 254
 <211> 560
 <212> DNA
 <213> Homo sapiens

<400> 254
 gaattcggca ccagtttggg gggtagaggt taattggaaa tggctctctg ggactgaaaa 60

```

ctgatgtttt tgcagattac ctcagggaaa cggaggtttg ttgagttaca gacacattaa 120
accaaaggcc gtgggaaaac ccctctccag ctccagggga ttggtcagga ccaccacta 180
accagtgcct tccttcttaa cattcacttt tagcagcttg tgtttatttt acatgggcag 240
ttttgatggg aaattgccat gaccacaggg gtttgaggtt ctgctttttt ttttcttct 300
tctttttcgg gggactgggg gactcctccc aagatcacat tttagcatct ttctctccta 360
ctccatttag aaaaataagt aacaggtgaa atgtgtctc agtggttaacg ggataattct 420
gctaccggct cctccctgat gattctgaaa tacactactg aacgagctct ggctggctct 480
ttctatcctg gatgtgggtt ttctgtgtag caattccttg atgtccagtt tggaaagatg 540
tactcttctc aacaagaaaa

```

```

<210> 255
<211> 612
<212> DNA
<213> Homo sapiens

```

```

<400> 255
gaattcggca ccaggcgggg cagcagggcc gcggccatgg ggagcttgaa ggaggagctg 60
ctcaaagcca tctggcacgc cttcacogac tgcaccagga ccacagggca aggtctccaa 120
gtcccagctc aaggctccttt ccataaacct gtgcacgggt ctgaagggtt ctcattgacct 180
agttgccctt gaagagcact tcagggatga tgatgagggt ccagtgtcca accagggcta 240
catgccttat ttaaacaggt tcatttttga aaaggctcaa gacaactttg acaagattga 300
attcaatagg atgtgttga ccctctgtgt caaaaaaaaa cctcaciaaag aatccccctgc 360
tcattacaga agaagatgca tttaaaatat gggttatttt caacttttta tctgaggaca 420
agtatccatt aattattgtg tcagaagaga ttgaatacct gcttaagaag cttacagaag 480
ctatgggagg aggttggcag caagaacaat ttgaacatta taaaatcaac tttgatgaca 540
gtaaaaatgg cctttctgca tgggaactta ttgagcttat tggaaatgga cagtttagca 600
aaggcatgga cc
612

```

```

<210> 256
<211> 1132
<212> DNA
<213> Homo sapiens

```

```

<400> 256
gaattcggca cgaggtctgg gagaggcctc tggagcagga ggcccagtg ctcttctgac 60
ccaaggcccc gcggtccagc ttctaagtgc cagatgatgg aggagcgtgc caacctgatg 120
cacatgatga aaactcagcat caagggtgtt ctccagtcgg ctctgagcct gggccgcagc 180
ctggatgcgg accatgcccc cttgcagcag ttctttgtag tgatggagca ctgcctcaaa 240
catgggctga aagttaagaa gagttttatt ggccaaaata aatcattctt tggctccttg 300
gagctggtgg agaaactttg tccagaagca tcagatatag cgactagtgt cagaaatctt 360
ccagaattaa agacagctgt ggggaagaggc cgagcgtggc tttatcttgc actcatgcaa 420
aagaaactgg cagattatct gaaagtgcct atagacaata aacatctctt aagcgagttc 480
tatgagcctg aggcctttaa gatggaggaa gaagggatgg tgattgttgg tctgctggtg 540
ggactcaatg ttctcgatgc caatctctgc ttgaaaggag aagacttggg ttctcaggtt 600
ggagtaatag atttttccct ctaccttaag gatgtgcagg atcttgatgg tggcaaggag 660
catgaaagaa ttactgatgt ccttgatcaa aaaaattatg tggagaagaa taaccggcac 720
ttgagctgca cagttgggga tcttcaaacc aagatagatg gcttggaata gactaactca 780
aagcttcaag aagagctttc agctgcaaca gaccgaattt gctcacttca agaagaacag 840
cagcagttaa gagaacaaaa tgaattaatt cgagaaagaa gtgaaaagag tgtagagata 900
acaaaacagg ataccaaagt tgagctggag acttacaagc aaactcggca aggtctggat 960
gaaatgtaca gtgatgtgtg gaagcagcta aaagaggaga agaaagtccg gttggaactg 1020
gaaaaagaac tggagttaca aattggaatg aaaaccgaaa tggaaattgc aatgaagtta 1080
ctggaagaag acaccacaga gaagcaggac acactagttg ccctccgcca gc 1132

```

```
<210> 260
<211> 994
<212> DNA
<213> Homo sapiens
```

gaattcggca	cgaggcggtt	cctgccttct	tgctgtctat	cagcctttct	tgcctcttcc	60
ttttcgctt	ccctgttctt	ccctttctca	aacaaacaag	acatggcaaa	ccgcagtcta	120
accagccct	ttgaaattat	ccatagtttt	acagacagct	ccaggccatg	agccacaatg	180
tccaaaatta	ttcttgagca	ctgatataaa	ttacttagac	cttctttgag	ggcagaactc	240
agctgttgt	ctcatgatgg	gcagtgtctg	aaagggttct	ggtatgtctt	caaaatgagt	300
ccacgagttt	actgagtgt	tacaggtaaa	ggaatgaata	taagatgtct	ttctgatcag	360
aacaggtgtc	ccttcacatg	agctttacta	gactctggga	gggaaaagta	gccaagtact	420
tctgaaccat	tttttaatac	ttgttttgct	atggtgaaat	tatagcagtt	atcccaaaat	480
gttttaatta	tcaaaatact	gtcttttaaa	aaaaaaaaaa	agtaaacacct	tttaaagcat	540
tagatttcac	ttgggtttct	tttccaaaaa	atgctaggtg	gacaaggcat	tgtaaacatg	600
agtttccttt	aagaaccatc	agaatataaa	tttaacatga	agaaaactgc	tatatctagt	660
agaaataata	tctaaagtgt	aacaactaaa	gtaccctcac	agaatagcaa	atacccttct	720
gttctggaca	tgggttcaaa	tttgaatatg	gaaataattt	ccttggaagt	ccctagaggc	780
aggtcagagg	aagtatgcat	taagagggaa	aggagagaat	ggaaataaaa	gtcactataa	840
tgcagattta	tgccttattt	tttagcattt	tttaaatgtt	gggtctttca	aggtgttttt	900
tgtcttttat	tagatctata	taaataagtt	aactagcaat	ttagttttgt	atttaagcta	960
cacttaatct	ttttcttttg	tqatatttat	ttct			994

<211> 594

<212> DNA

<213> Homo sapiens

 $\langle 220 \rangle$

<221> misc feature

 $\langle 222 \rangle \quad (538)^{-}$

<223> n=A, T, C or G

gaattcggca	ccagtggaga	tccagctgaa	ccatgccaac	cgccaggctg	cggaggcaat	60
caggaaacctt	cggaacaccc	agggaaatgct	gaaggacaca	cagctgcacc	tggacgatgc	120
tctcagaggc	caggacgacc	tgaaagagca	gctggccatg	gttgagcgca	gagccaacct	180
gatgcaggct	gagatcgagg	agctcagggc	atccctggaa	cagacagaga	ggagcaggag	240
agtggccgag	caagagctac	tggatgccag	tgagcgcgtg	cagctcctcc	acaccagaa	300
caccagcctc	atcaacacca	agaagaagct	ggagacagac	atttcccaaa	tccagggaga	360
gatggaagac	atcgtccagg	aagcccgcaa	cgcagaagag	aaggccaaga	aagccatcac	420
tgatgccgcc	atgatggcgg	aggagctgaa	gaaggagcag	gacaccagcg	cccacctgga	480
gcggatgaag	aagaacatgg	agcagaccgt	gaaggacctg	cagcaccgtc	tggacgangc	540
tgagcagctt	ggcgctgaag	ggcgggcaag	aagcagatcc	agaaactgga	ggct	594

<211> 594

<212> DNA

<213> Homo sapiens

gaaaaggttg	ctggagccaa	aggcatagtc	agggttaatg	ctcctttttc	tttatcccaa	60
atcagatagt	gtttaggctt	tttcatcaaa	tataaaaacc	cagcccagtt	catggctcat	120
tcggcagcaa	ccctgagacg	ctttacagct	ctagacccta	aaaggtcaaa	aggcgcgtct	180
atgctcaata	tacattttat	tacccaatct	gccccggaca	ttaataaaaa	ctccaaaaaat	240
taaatccggc	cctcaaaccc	cacaacagga	cttaattgac	ctcaccttca	agggtgtagaa	300
taataaaaaa	aaaaagttgc	aattccttgc	ctccgctgtg	agacaaaacc	cagccacatc	360


```

gccttgggat atggtttttt cctggagttg aacttggcag aagctctcaa gttcattgat 360
cgtaagagct ctctcctcac agagctcagc aacagcctca ccaaggactc catgaatatc 420
aaagcccata tccacatgtt gctagagggg cttagagaaac tacaaggcct gcagaatttc 480
ccagagaagc ctcaccattg acttcttccc cccatcctca gacattaaag agcc 534

```

<210> 266

<211> 552

<212> DNA

<213> Homo sapiens

<400> 266

```

gaattcggca ccagggcacc tccgcctcgc cgccgctagg tcggccgggt ccgcccgggt 60
gccgcctagg atgaatatca tggacttcaa cgtgaagaag ctggcggccg acgcaggcac 120
cttcctcagt cgcgcctgtc agttcacaga agaaaagctt ggccaggctg agaagacaga 180
attggatgct cacttagaga acctccttag caaagctgaa tgtaccaaaa tatggacaga 240
aaaaataatg aaacaaactg aagtgttatt gcagccaaat ccaaagtcca ggatagaaga 300
atttgtttat gagaaactgg atagaaaagc tccaagtcgt ataaacaacc cagaactttt 360
gggacaatat atgattgatg cagggactga gtttgccca ggaacagctt atggtaatgc 420
ccttattaaa tgtggagaaa ccaaaaaaag aattggaaca gcagacagag aactgattca 480
aacgtcagcc ttaaattttc ttactccttt aagaaacttt atagaaggag attacaaaac 540
aattgctaaa ga 552

```

<210> 267

<211> 551

<212> DNA

<213> Homo sapiens

<400> 267

```

gaagcctacc agccagggtgc cgccccccc acccccggcc cagccccctc ctgcagcggt 60
ggaagcgggt cggcagatcg agcgtgaggg ccagcagcag cagcacctgt accgggtgaa 120
catcaacaac agcatgcccc caggacgcac gggcatgggg accccgggga gccagatggc 180
ccccgtgagc ctgaatgtgc cccgacccaa ccaggtgagc gggcccgtca tgcccagcat 240
gcctcccggg cagtggcagc aggcgccccct tccccagcag cagcccatgc caggcttgcc 300
caggcctgtg atatccatgc agggccaggc ggccgtgggt gggccccgga tgcccagcgt 360
gcagccaccc aggagcatct caccagcgc tctgcaagac ctgctgcgga ccctgaagtc 420
gccagctcc cctcagcagc aacagcaggt gctgaacatt ctcaaataca acccgagct 480
aatggcagct ttcatacaac agcgcacagc caagtacgtg gccaatcagc ccggcatgca 540
gccccagcct g 551

```

<210> 268

<211> 573

<212> DNA

<213> Homo sapiens

<400> 268

```

gaattcggca ccagggttcc ttgtgggcta gaagaatcct gcaaaaatgt ctctctatcc 60
atctctcgaa gacttgaagg tagacaaagt aattcaggct caaactgctt tttctgcaaa 120
ccctgccaat ccagcaattt tgtcagaagc ttctgtcctt atccctcacg atggaaatct 180
ctatcccaga ctgtatccag agctctctca atacatgggg ctgagtttaa atgaagaaga 240
aatacgtgca aatgtggccg tggtttcttg tgcaccactt caggggcagt tggtagcaag 300
accttccagt ataaactata tgggtggctcc tgtaactggt aatgatgttg gaattcgtag 360
agcagaaaatt aagcaaggga ttcgtgaagt catttttgtt aaggatcaag atggaaaaat 420
tggtactcag cttaaatcaa tagataatgg tatatttgtt cagctagtcc aggttaattc 480
tccagcctca ttggttgggtc tgagatttgg ggaccaagta cttcagatca atggtgaaaa 540

```


<213> Homo sapiens

<400> 284

```
gaattcggca cgaggcggag gccgcggagg ctctcgggtc cttcagcacc cctcggccccg 60
acgcacccac gcccctcacc ccccagagagc cgaaaatgga cccaagtggg gtcaaagtgc 120
tgaaaacagc agaggacatc caggagaggg ggcagcaggt cctagaccga taccaccgct 180
tcaaggaact ctcaaccctt aggcgtcaga agctggaaga ttcctatcga ttccagttct 240
ttcaaagaga tgctgaagag ctggagaaat ggatacagga aaaacttcag attgcatctg 300
atgagaatta taaagaccca accaacttgc agggaaagct tcagaagcat caagcatttg 360
aagctgaagt gcaggccaac tcaggagcca ttgttaagct ggatgaaact ggaaacctga 420
tgatctcaga agggcatttt gcatctgaaa ccatacggac ccgtttgatg gagctgcacc 480
gccagtggga attacttttg gagaagatgc gaga 514
```

<210> 285

<211> 383

<212> DNA

<213> Homo sapiens

<400> 285

```
gaattcggca cgaggccggg ctccacccgg catcctgctc cactctggcg accgcccccg 60
gggccccgc gcggggcgcg gcgcccgcga tgggcgagga ggactactat ctggagctgt 120
gcgagcggcc ggtgcagttc gagaaggcga accctgtcaa ctgcgtcttc ttcgatgagg 180
ccaacaagca ggtttttgct gttcgatctg gtggagctac tggcgtggta gttaaaggcc 240
cagatgatag gaatcccatc tcatttagaa tggatgacaa aggagaagtg aagtgcatta 300
agttttcctt agaaaataag atattggctg ttcagaggac ctcaaagact gtggattttt 360
gtaattttat cctgataat tcc 383
```

<210> 286

<211> 943

<212> DNA

<213> Homo sapiens

<400> 286

```
gaattcggca ccagggccgt ggcgaggag gagcgctgca cgggtggagcg tcgggcccgc 60
ctcacctacg cggagttcgt gcagcagtac gtgcgcccct gatcgcgag gtgcgctcct 120
gttcaccggc ccgtctgccc cgaccgcccga aggcgcgctt cccctgacct cgcgcgcacg 180
cgtggggctg gggcggcgag gctggcggtc cggcctggcc gcgactctgc cttcttttcc 240
agaggttccg ggccctgtgc tcccgcgaca ggttgctggc ttcgtttggg gacagagtgg 300
tccggctgag caccgccaac acctactcct accacaaagt ggacttgccc ttccaggagt 360
atgtggagca gctgctgcac ccccaggacc ccacctcctt gggcaatggg gaggcagccc 420
taggcggcgg tagggggtgg ggacgcttgg agtctccagg tgccaggatc cctgtccccg 480
ccgtctctgt tggcagacac cctgtacttc ttcggggaca acaacttcac cgagtgggccc 540
tctctcttcc ggcactactc cccacccccca tttggcctgc tgggaaccgc tccagcttac 600
agctttggaa tcgcaggagc tggctcgggg gtgcccttcc actggcatgg acccggttac 660
tcagaagtga tctacggtcg taagcgttgg ttcctttacc cacctgagaa gacgccagag 720
ttccacccca acaagaccac actggccttg ctccgggaca catacccagc cctgccaccg 780
tctgcacggc ccctggagtg taccatccgg gctggtgagg tgctgtactt ccccgaccgc 840
tggtggcatg ctacgctcaa ccttgacacc agcgtcttca tctccacctt cctcggctag 900
ccaaaacagc tggcaggact gccggtcaca caccagcacg tcc 943
```

<210> 287

<211> 1143

<212> DNA

<213> Homo sapiens

<400> 287

```

gaattcggca cgagggaaga acagctgttg gaacaacaag aatattttaga aaaagaaatg 60
gaggaagcaa agaaaatgat atcaggacta caggccttac tgctcaatgg atccttacct 120
gaagatgaac aggagaggcc cttggccctc tgtgaaccag gtgtcaatcc cgaggaacaa 180
ctgattataa tccaaagtcg tctggatcag agtatggagg agaatacagga cttaaagaag 240
gaactgctga aatgtaaaca agaagccaga aacttacagg ggataaagga tgccttgag 300
cagagattga ctacgcagga cacatctgtt cttcagctca aacaagagct actgagggca 360
aatatggaca aagatgagct gcacaaccag aatgtggatc tgcagaggaa gctagatgag 420
aggaaccggc tcttgggaga atataaaaaa gagctggggc agaaggatcg ccttcttcag 480
cagcaccagg ccaagttaga agaagcactc cggaaactct ctgatgtcag ttaccaccag 540
gtggatctag agcgagagct agaacacaaa gatgtcctct tggctcactg tatgaaaaga 600
gaggcagatg aggcgaccaa ctacaacagt cacaactctc aaagcaatgg ttttctcctt 660
ccaacggcag gaaaaggagc tacttcagtc agcaacagag ggaccagcga cctgcagctt 720
gttcgagatg ctctccgcag cctgcgcaac agcttcagtg gccacgatcc tcagcaccac 780
actattgaca gcttggagca gggcatttct agcctcatgg agcgctgca tgttatggag 840
acgcagaaga aacaagaaag aaaggttcgg gtcaagtcac ccagaactca agtaggtagt 900
gaataccggg agtcctggcc ccctaactca aagttgcctc actcacagag ctctccaact 960
gtcagcagca cctgtactaa agtgccttat ttcactgacc ggtcacttac gcccttcagt 1020
gtcaatatac caaagagggt ggaggagggt acgttaaagg attttaaagc agctattgat 1080
cggaaggaa atcaccggta tcacttcaaa gcactggatc ctgagtttgg cactgtcaaa 1140
gag 1143

```

<210> 288

<211> 881

<212> DNA

<213> Homo sapiens

<400> 288

```

gtgagagcgg gccgaggaga ttggcgacgg tgtcgcccg gttttcgttg gcgggtgcct 60
gggctggttg gaacagccgc ccgaagggaag caccatgatt tcggccgcgc agttgttgga 120
tgagttaatg ggccgggacc gaaacctagc cccggacgag aagcgcagca acgtgcggtg 180
ggaccacgag agcgtttgta aatattatct ctgtggtttt tgtcctgcgg aattgttcac 240
aaatacacgt tctgatcttg gtccgtgtga aaaaattcat gatgaaaatc tacgaaaaca 300
gtatgagaag agctctcggt tcatgaaagt tggctatgag agagattttt tgcgatactt 360
acagagctta cttgcagaag tagaacgtag gatcagacga ggccatgctc gtttggcatt 420
atctcaaaac cagcagtcct ctggggccgc tggcccaaca ggcaaaaatg aagaaaaaat 480
tcaggttcta acagacaaaa ttgatgtact tctgcaacag attgaagaat tagggctctga 540
aggaaaagta gaagaagccc aggggatgat gaaattagtt gagcaattaa aagaagagag 600
agaactgcta aggtccacaa cgtcgacaat tgaaagcttt gctgcacaag aaaaacaaat 660
ggaagtttgt gaagtatgtg gagccttttt aatagtagga gatgccagc cccgggtaga 720
tgaccatttg atgggaaaac aacacatggg ctatgccaaa attaaagcta ctgtagaaga 780
attaaaagaa aagttaagga aaagaaccga agaactgat cgtgatgagc gtctaaaaaa 840
ggagaagcaa gaaagagaaa aaaaaaaaaa aaaaactcga g 881

```

<210> 289

<211> 987

<212> DNA

<213> Homo sapiens

<400> 289

```

gaattcggca cgagggactg tggtttccag gaatggtggc gtctcacgct tcttgtgctt 60
tttcttttgg ggctccgag cggtgggggt tgggggactg ggcaggaggc tccctgtaaa 120
catttggact tgggctgggg caggggctgg tgttgggcaa agctgggggt ccaggctgga 180

```



```

caggctgact gtgctgtcct gattgttgc tctggtgttg gtgaatttga agctggtatc 420
tccaagaatg ggcaggaccc gagagcatgc ctttctggct tacacactgg gtgtgaaaca 480
actaattgtc ggtgttaaca aaatggatt                                     509

```

```

<210> 298
<211> 267
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(267)
<223> n = A,T,C or G

```

```

<400> 298
gggacggggg aaaggagacg cttcttcctc ttgctgctct tctcgttccc gagatcagcg 60
gcggcgggtga ccgcgagtgg gtcggcaccg tctccggctc cggngcnaa caatgctgac 120
tgatagcgga ggcggnggca cctccttnna ggaggacctg gactctgtgg ctccgcgac 180
cgccccagct ggggcctcgg agccgcctcc gccgggaggg gtcggtctgg ggatccncac 240
cgngaggctn tttggggagg gcgggcc                                     267

```

```

<210> 299
<211> 121
<212> DNA
<213> Homo sapien

```

```

<400> 299
ggcacgaggg ccctcggagc tcgtttccag atcgaggtaa gagggacttt cttaaaggcc 60
tagtctatgg gatggggcgg cggagggaat tttttgagaa ataaaaatgaa gctgcagtgt 120
a                                     121

```

```

<210> 300
<211> 533
<212> DNA
<213> Homo sapien

```

```

<400> 300
aaggtgcaca gtatttgatg caggctgctg gtcttggtcg tatgaagcca aacacacttg 60
tccttggtatt taagaaagat tggttgcaag cagatatgag ggatgtggat atgtatataa 120
acttatttca tgatgctttt gacatacaat atggagtagt ggttattcgc ctaaaagaag 180
gtctggatat atctcatctt caaggacaag aagaattatt gtcatcaca gagaaatctc 240
ctggcaccaa ggatgtggtg gtaagtgtgg aatatagtaa aaagtccgat ttagatactt 300
ccaaaccact cagtgaaaaa ccaattacac acaaagttga ggaagaggat ggcaagactg 360
caactcaacc actgttgaaa aaagaatcca aaggccctat tgtgccttta aatgtagctg 420
acaaaaagct tcttgaagct agtacacagt ttcagaaaaa acaaggaaag aatactattg 480
atgtctgggtg gctttttgat gatggaggtt tgaccttatt gataccttac ctt       533

```

```

<210> 301
<211> 560
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature

```


<210> 304
 <211> 441
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(441)
 <223> n = A,T,C or G

<400> 304
 gctggacgga gacctgctgg aggaggagga gctggaggaa gcagaggagg aggaccggtc 60
 gtcgtgctg ctgctgtgc cgccgcggc caccgcctct cagaccagc agatcccagg 120
 cgggtccctg gggctctgtc tgctgccagc cgccagggtc gatgccggg aggcggcggc 180
 ggcggcgggg gtgctgtacg gaggggacga tgcccagggc atgatggcgg cgatgctgtc 240
 ccacgcctac ggccccggcg gttgtggggc ggccggcgcc gccctgaacg gggagcaggc 300
 ggccctgtc cggagaaaga gcgtcaacac caccgagtgc gtcccggtgc ccagctccga 360
 gcacgtcgcc gagatcgctg gccgccaggg ttgtaaaatt aaagcactga nagccaagac 420
 aaacacgtat atcaagactc c 441

<210> 305
 <211> 491
 <212> DNA
 <213> Homo sapien

<400> 305
 tcgccatgcc cccttcttag cactgcaccg ccagggtccat gctgctgcca cccagacct 60
 gggctttgcc tgccacctct gtgggcagag cttccgaggc tgggtggccc tggttctgca 120
 tctgcgggcc cattcagctg caaagcggcc catcgcttgt cccaaatgcg agagacgctt 180
 ctggcgacga aagcagcttc gagctcatct gcggcggtgc caccctcccg ccccgagggc 240
 ccggcccttc atatgcggca actgtggccg gagctttgcc cagtgggacc agctagtgtc 300
 ccacaagcgg gtgcacgtag ctgaggccct ggaggaggcc gcagccaagg ctctggggcc 360
 ccggcccagg ggccgccccg cggtgaccgc cccccggccc ggtggagatg ccgtcgaccg 420
 ccccttccag tgtgcctgtt gtggcaagcg cttccggcac aagcccaact tgatcgctca 480
 cccgcgcgtg c 491

<210> 306
 <211> 547
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(547)
 <223> n = A,T,C or G

<400> 306
 tctctttctt ttaagacagg aatgtaagcc acaacattta caaatacaat gttttaactc 60
 tctacatgta ggaagccaac ctgctccttt ttgatcttct tctttggcac aacctcagtg 120
 gatttctctg attcagaacg agttctaatt gatcttctct gttgcttctt ttctactgag 180
 cctgtagaac cagatgttgc ttcaggagat gatacactct gcgttggctt ttcatttctc 240
 tggtttggtg tagaaattat aagcctgtct tgccccctga cacttatttc tgttttgta 300
 ccaattccct ttgttgaata aacaaattga tcgataaatt tcccatcccc tgtagcattc 360
 tgaagagcaa acacttggtc aattttcaca actggagaca tgttacactt ctgcaaatcc 420

aggctccctt tgtgcatccg taatggaagc tggtaaggat ttccttgctg ccgcagtttt 480
 ccaggctatt ttaacaggcg gnggctcttc ctctttccgc acttggtgtgc cgcctctggc 540
 tatgtct 547

<210> 307
 <211> 571
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(571)
 <223> n = A,T,C or G

<400> 307
 cgctgcatgt gataatgtca tcatttattt ttaaatgggt ctaaattgca natttaagtt 60
 gatttcaaat caaccctatt tttaaattac ttttaatagg aanaaatgaa gcaaggacat 120
 acataatcta ctatatattga aggactcaaa caaatacatg tttggctgtg aattctgtac 180
 tctcaccaaa acagagataa aaatccacct aaaatacact ttccttcatt tagtgcttgt 240
 ggganaaggt caagtattgc actttaaaat tactttcatc taacatttgc cccaactttc 300
 cccctgaatt cactatatgt tttcagcaaa catgatttta taaattttaa gtataaaaagc 360
 aactaggttt tctaattcaa ctttggaagg ttacttttac tctacanagc tatttttgta 420
 aaacggcata ttactttaca aaattganag ataggggcat ccagctgagg tacatttcct 480
 cccttggcgt tgagtttctg gacttgggtc gggggcacag gcttggtgtga ctgccccgtg 540
 gcccgataca tggcctggac cccaggatgc g 571

<210> 308
 <211> 591
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(591)
 <223> n = A,T,C or G

<400> 308
 ctcttatgt gtctgcctac ttcatttttc ggcatttccct gcttatccaa gttcaccatt 60
 tcaggtcacc actggatc agttgcctgt atataattat caggcatttc ctgcttatcc 120
 aagttcacca tttcagggtca ccaactggata tcagttgcct gtatataatt atcaggcatt 180
 tctgcttat ccaagttcac catttcagggt caccactgga tatcagttgc ctgtatataa 240
 ttatcaggca tttcctgctt atocaaagttc accatttcag gtcaccactg gatatcagtt 300
 gcctgtatat aattatcagg catttccctgc ttatccaagt tcaccatttc aggtcaccac 360
 tggatatcag ttgcctgtat ataattatca ggcatttccct gcttatccaa gttcaccatt 420
 tcaggtcacc actggatc agttgcctgt atataattat caggcatttc ctgcttatcc 480
 aaattcagca gttcagggtca ccaactggata tcagttccat gtatacaatt accagatgcc 540
 accgcagtgc cctgttgggg gagcaaagga gaaatntgtg gaccgaagca t 591

<210> 309
 <211> 591
 <212> DNA
 <213> Homo sapien

<400> 309

<220>
 <221> misc_feature
 <222> (1)...(591)
 <223> n = A,T,C or G

<400> 318
 gatggcgtagc ttggcttggg gactggcgcg cgcttcgtgt ccgagttctc tgcaggtcac 60
 tagtttcccg gtagttcagc tgcacatgaa tagaacagca atgagagcca gtcagaagga 120
 ctttgaaaat tcaatgaatc aagtgaact cttgaaaaag gatccaggaa acgaagtga 180
 gctaaaactc tacgcgctat ataagcaggg cactgaagga ccttgtaaca tgcccaaacc 240
 aggtgtattt gacttgatca acaaggccaa atgggacgca tggaatgccc ttggcagcct 300
 gcccaaggaa gctgccaggc agaactatgt ggatttgggtg tccagtttga gtccttcatt 360
 ggaatcctct agtcaggtgg agcctggaac agacaggaaa tcaactgggt ttgaaactct 420
 ggtggtgacc tccgaagatg gcatcacaaa gatcatgttc aaccggccca aaaagaaaaa 480
 tgccataaac actgagatgt atcatgaaat tatgcgtgca cttaaagctg ccagcaanga 540
 tgactcaatc atcacttgtt ttaacaggaa atggtgacta ttacagtagn g 591

<210> 319
 <211> 591
 <212> DNA
 <213> Homo sapien

<400> 319
 gaattcggga cgaggttgct gctaagcgaa cgccctttgg agcttacgga ggcctttctga 60
 aagacttcac tgctactgac ttgtctgaat ttgctgccaa ggctgccttg tctgctggca 120
 aagtctcacc tgaaacagtt gacagtgtga ttatgggcaa tgcctgcag agttcttcag 180
 atgctatata tttggcaagg catgttggtt tgcgtgtggg aatcccaaag gagacccag 240
 ctctcacgat taataggctc tgtggttctg gttttcagtc cattgtgaat ggatgtcagg 300
 aaatttgtgt taaagaagct gaagttgttt tatgtggagg aaccgaaagc atgagccaag 360
 ctccctactg tgtcagaaat gtgcgttttg gaaccaagct tggatcagat atcaagctgg 420
 aagattcttt atgggtatca ttaacagatc agcatgtcca gctcccatg gcaatgactg 480
 cagagaatct tgctgtaaaa cacaaaataa gcagagaaga atgtgacaaa tatgccctgc 540
 agtcacagca gagatggaaa gctgctaatt atgctggcta ctttaatgat g 591

<210> 320
 <211> 591
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(591)
 <223> n = A,T,C or G

<400> 320
 ggctccggcg tctgcagggg tgcgcgagct aaccctgtggc taggcgagtg gggcggggcg 60
 gccggcacca tgtcgaggca ggcgaaccgt ggcaccgaga gcaagaaaat gagctctgag 120
 ctcttcaccc tgacctatgg tgccctgggtc acccagctat gtaaggacta tgaaaatgat 180
 gaagatgtga ataaacagct ggacaaaatg ggctttaaca ttggagtcgg gctgattgaa 240
 gatttcttgg ctcggtcaaa tgttgggagg tgccatgact ttcgggaaac tgcggatgtc 300
 attgccaaagg tggcgttcaa gatgtacttg ggcatactc caagcattac taattggagc 360
 ccagctgggtg atgaattctc cctcattttg gaaaataacc ccttggtgga ctttgtggaa 420
 cttctgata accactcatc ccttattttat tccaatctct tgtgtggggg gttgcgggga 480

ttggcatttt gtaaagacga ccoctgcagnc ccoctgtttgn aactttttta ataaaaataga 480
catctgttta cttg 494

<210> 332
<211> 538
<212> DNA
<213> Homo sapien

<400> 332
aaagaacaaa tggaaacgcga tgggtgtttct gaacaagagt ctcaaccgtg tgcattttatt 60
gggataggaa atagtgacca agaaatgcag cagctaaact tggaggaaa gaactattgc 120
acagccaaaa cattgtatat atctgactca gacaagcgaa agcacttcat gttgtctgta 180
aagatgttct atggcaacag tgatgacatt ggtgtgttcc tcagcaagcg gataaaaagtc 240
atctccaaac cttccaaaaa gaagcagtc tggaaaaatg ctgacttatg cattgcctca 300
ggaacaaagg tggctctgtt taatcgacta cgatcccaga cagttagtac cagatacttg 360
catgtagaag gaggtaatat tcatgccagt tcacagcagt ggggagcctt tttattcat 420
ctcttgatg atgatgaatc agaaggagaa gaattcacag tccgagatgg ctacatccat 480
tatggacaaa cagtcaaact tgtgtgctca gttactggca tggcactccc aagattga 538

<210> 333
<211> 499
<212> DNA
<213> Homo sapien

<400> 333
ctcagcctgc gggactgctc ggctcggett ctaggcggtt ttgatgaaca cctggcttta 60
ttcttgcaat gaagaaagg tctcaacaaa aaatattctc caaagcaaag ataccatcat 120
catctcactc tcctatccca tcatctatgt ccaatatgag atctaggtca ctttcacctt 180
tgattggatc agagactcta ctttttcatt ctggaggaca gtggtgtgag caagttgaga 240
ttgcagatga aaacaatatg cttttggact atcaagacca taaaggagct gattcacatg 300
caggagttag atatatata gagggcctca ttaaaaaact tactaaacag gataatttgg 360
ctttgataaa atctctgaac ctttcacttt ctaaagacgg tggcaagaaa ttttaagtata 420
ttgagaattt ggaaaaatgt gttaaacttg aagtactgaa tctcagctat aatctaatag 480
ggaagattga aaagtgcga 499

<210> 334
<211> 561
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(561)
<223> n = A,T,C or G

<400> 334
ttcccggtag ttcagctgca catgaataga acagcaatga gagccagtca gaaggacttt 60
gaaaattcaa tgaatcaagt gaaactcttg aaaaaggatc caggaaacga agtgaagcta 120
aaactctacg cgctatataa gcaggccact gaaggacctt gtaacatgcc caaaccaggt 180
gtatttgact tgatcaacaa ggccaaatgg gacgcatgga atgcccttgg cagcctgcc 240
aaggaagctg ccaggcagaa ctatgtggat ttggtgtcca gtttgagtcc ttcattggaa 300
tcctctagtc aggtggagcc tggaaacagac aggaaatcaa ctgggtttga aactctggtg 360
gtgacctccg aagatggcat cacaaagatc atgttcaacc cggcccaaaa agaaaaatgc 420
cataaacact gagatgtatc atgaaattat gcgtgcactt aaagctgcc aagagatga 480

ctcaatcatc actgttttaa cangaaatgg tgactattac agtagtgga atgatctgac 540
taacttcnct gatattcccc c 561

<210> 335
<211> 551
<212> DNA
<213> Homo sapien

<400> 335
aagctggtca tggctgggga gaccaccaac tcccgcggcc agcggctgcc ccagaaggga 60
gacgtggaga tgctgtgagg cgggcccggc tgccagggct tcagcggcat gaaccgcttc 120
aattcgcgca cctactccaa gttcaaaaac tctctggtgg tttccttcct cagctactgc 180
gactactacc ggccccgggt cttctctctg gagaatgtca ggaactttgt ctcttcaag 240
cgctccatgg tectgaagct caccctccgc tgcctggtcc gcatgggcta tcagtgcacc 300
ttcggcgtgc tgcaggccgg tcagtacggc gtggcccaga ctaggaggcg ggccatcatc 360
ctggccggcg cccctggaga gaagctccct ctgttcccgg agccactgca cgtgtttgct 420
ccccgggcct gccagctgag cgtggtgggt ggatgacaag aagtttgtga gcaacataac 480
caggttgagc tcgggtcctt tccggacat acggtgcgag aaacgatgtc cgacctgccg 540
gaagtgcgga a 551

<210> 336
<211> 540
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(540)
<223> n = A,T,C or G

<400> 336
aggtctatgt ctactgaagg caataaacga ggaatgatcc agcttattgt tgcaaggaga 60
ataagcaagt gcaatgagct gaagtcacct gggagccccc ctggacctga gctgcccatt 120
gaaacagcgt tggatgatag agaacgaaga atttccatt ccctctacag tgggattgag 180
gggcttgatg aatcgcccag cagaaatgct gccctcagta ggataatggg taaataaccag 240
ctgtccccta cagtgaatat gcccacagat gacactgtca ttatagaaga tgacaggttg 300
ccagtgttc ctccacatct ctctgaccag tctcttcca gctcccatga tgatgtgggg 360
tttgtgacgg cagatgctgg tacttgggccc aaggctgcaa tcagtgttc agccgactgc 420
tctttgagtc cagatgttga tccagttctt gcttttcaac gaaaaaggat ttggacgtca 480
gaagtatgtc agaaaaacgc accaaagcaa ttttcanatg ccagtcaatt ggatttcgtt 540

<210> 337
<211> 422
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(422)
<223> n = A,T,C or G

<400> 337
gcagcaggaa cagttacagc agcagcagca acagcagctg ttgcaacagc agcaggaaca 60
attgcagcag caacaactgc agcctcctcc cctggagccc gaggaggagg aagaggtgga 120

```

gctggagctc atgccggtgg acctggggtc agagcaggag ctggagcagc agcggcagga 180
gttggagcgg cagcaggagc tggaacggca gcaggagcag cggcagctgc agctcaaact 240
gcaggaggag ctgcagcagc tggagcaaca gctggagcag cagcagcagc agctggagca 300
gcaggagggtg cagctggagc tgaccccggt ggagctaggc gccagcagc aggagggtgca 360
gctggagctg acccccgctgc agccggagct gcagctggaa ctggtgccan cccagggggc 420
gg                                     422

```

```

<210> 338
<211> 601
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(601)
<223> n = A,T,C or G

```

```

<400> 338
catcttacga acgctctatg atgtcttatg agcgggtctat gatgtccctt atggttgaac 60
gctctatgat gtcagcctac gagcgtctta tgatgtcagc ctacgagcgc tctatgatgt 120
cccctatggc tgagcgtctt atgatgtcag cttatgaacg ctccatgatg tcagcttatg 180
aacgctccat gatgtcccca atggctgacg gatctatgat gtccatgggt gctgaccggg 240
ctatgatgtc gtcatactct gctgctgacc ggtctatgat gtcacgtac tctgcagctg 300
accgatctat gatgtcatct tatactgctg atcgttcaat gatgtctatg gctgctgatt 360
cttacaccga ttcttacact gacacatata cagaggcata tatggtgcc a ctttgcctc 420
ctgaagagcc cccaacaatg ccaccgttgc cacctgagga gccaccaatg acaccaccat 480
tgccctnctga ggaaccaccc agagggtcca gcattgccca cttgagcagt cagcattaac 540
cagcttgaaa atacttggcc ctacanangg tgccatcatt accatctgaa gagctgtatc 600
g                                     601

```

```

<210> 339
<211> 440
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(440)
<223> n = A,T,C or G

```

```

<400> 339
agagggagga ggcccaactg gtgatgctgc tgctgctgct gctgccgccg ccgccgcctc 60
tattgctgat actctagtgg ggctggaagg gtggttccta ttgcgaccat cgccaaccag 120
agacagaggg aaaaaaaaaa ccggcagcca ctgctgatgt tgggttcgga ggctgcatcc 180
gactcgggtca caaggaaaat ggattcagtt tgcatctctc cctcctttta acagcttctc 240
cgggtctcag catggtatca aagcttgaaa gagagaagac tcaagaagcg aagaggattc 300
gtgagctgga gcagcgcaag cacacggtgc tgggtgacaga actcaaagcc aagctccatg 360
aggagaagat gaaggagctg caggctgtga gggagaacct tatcaagcag cacgacagga 420
aatgtcaang acggtgaagg                                     440

```

```

<210> 340
<211> 450
<212> DNA
<213> Homo sapien

```

<220>
 <221> misc_feature
 <222> (1)...(450)
 <223> n = A,T,C or G

<400> 340
 gatttccagg ggcggatatt gagtgtcgac ccagaggaag aaagggagga gggcccgcct 60
 aggattcctc aggccgacca gtggaagtct tcaaacaaga gcctgggtgga ggctctgggg 120
 ctggaagccg aggggtgcagt tcctgagaca cagactttga ccggatggag taaggggttc 180
 attggcatgc acagggaaat gcaagtcaac cccatttcaa agcggatggg gcccatgact 240
 gtggtcagga tggacgttcc agtccagcca ggcccttttc ggaccctgct ccagtttctt 300
 tatacgggac aactggatga aaagggaaaag gatttggtgg gcctggctca gatcgagag 360
 gtcctcgaga tgttcgattt gaggatgatg gtggaaaaca tcatgaacaa ggaagccttc 420
 atgaaccagg agattacgaa nncctttcac 450

<210> 341
 <211> 451
 <212> DNA
 <213> Homo sapien

<400> 341
 aacagctatt aaaacagaaa atggatgaac ttcataagaa gttgcatcag gtgggtggaga 60
 catcccatga ggatctgccc gcttcccagg aaaggtccga ggtaaatcca gcacgtatgg 120
 ggccaagtgt aggcctcccag caggaactga gagcgccatg tcttccagta acctatcagc 180
 agacaccagt gaacatggaa aagaacccaa gagaggcacc tcctgttgtt cctccttttg 240
 caaatgctat ttctgcagct ttgggtgtccc cagccaccag ccagagcatt gtcctcctg 300
 ttccttttgaa agcccagaca gtaacagact ccatgtttgc agtggccagc aaagatgctg 360
 gatgtgtgaa taagagtact catgaattca agccacagag tggagcagag atcaaagaag 420
 ggtgtgaaac acataagggt gccaacacaa g 451

<210> 342
 <211> 498
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(498)
 <223> n = A,T,C or G

<400> 342
 ctcaagcagg ctattgaaga ggaaggaggc gatccagata atattgaatt aactgtttca 60
 actgatactc caaacaagaa accaactaaa ggcaaaggta aaaaacatga agcagatgag 120
 ttgagtggag atgcttctgt gggaagatga tgcttttatc aaggactgtg aattggagaa 180
 tcaagaggca catgagcaag atggaaatga tgaactaaag gactctgaag aatttggtga 240
 aaatgaagaa gaaaatgtgc attccaagga gttactctct gcagaagaaa acaagagagc 300
 tcatgaatta atagaggcag aaggaataga agatatagaa aaagaggaca tcgaaagtca 360
 ggaaattgaa gctcaagaag gtgaagatga tacctttcta acagcccaag atggtgagga 420
 agaagaaaat gagaaagata tagcagggtt ctggtgatgg cncacaagaa gtatntaaac 480
 ctcttccttc aaaaagggt 498

<210> 343
 <211> 491

<212> DNA
<213> Homo sapien

<400> 343
ccgaccccta ctcggcggcg caactccaca accagtagcg ccccatgaat atgaacatgg 60
gtatgaacat ggcagcagcc gcggcccacc accaccacca ccaccaccac caccocggtg 120
cctttttccg ctatatgcgg cagcagtgca tcaagcagga gctaattctgc aagtggatcg 180
accccgagca actgagcaat cccaagaaga gctgcaacaa aactttcagc accatgcacg 240
agctggtgac acacgtctcg gtggagcacg tcggcgggcc ggagcagagc aaccacgtct 300
gcttctggga ggagtgtccg cgcgagggca agcccttcaa ggccaaatac aaactggtea 360
accacatccg cgtgcacaca ggcgagaaac ccttcctctgc cttccgggt gtggcaaagt 420
cttcgcgcgc tccgagaacc tcaagatcca caaaaggacc acacagggga gaagccgtcc 480
agtggagttg a 491

<210> 344
<211> 412
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(412)
<223> n = A,T,C or G

<400> 344
gtgcgctgtc ttcccgcttg cgtcagggac ctgcccgaact cagtggccgc catggcatca 60
gatgaaggca aactttttgt tggagggctg agttttgaca ccaatgagca gtcgctggag 120
caggtcttct caaagtacgg acagatctct gaagtgggtg ttgtgaaaga cagggagacc 180
cagagatctc ggggatttgg gtttgtcacc tttgagaaca ttgacgacgc taaggatgcc 240
atgatggcca tgaatgggaa gtctgtagat ggacggcgaga tccgagtaga ccaggcaggc 300
aagtcgtcan acaaccgatc ccgtgggtac cgtgggtggct ctgccggggg ccgggggcttc 360
ttccgtgggg gcccgangac ggggcccgtg ggttctctaa aagaagaggg ga 412

<210> 345
<211> 498
<212> DNA
<213> Homo sapien

<400> 345
aactagtctc gggccatcct ttctgcgcac ccggtgtcgc tgggctgcac cccgggcggg 60
gacgtccgcc gggcacggga gggggccaag atgccgatca ataaatcaga gaagccagaa 120
agctgcgata atgtgaaggt tgttgtagg tgccggcccc tcaatgagag agagaaatca 180
atgtgctaca aacaggctgt cagtgtggat gagatgaggg gaactatcac tgtacataag 240
actgattctt ccaatgaacc tccaaagaca tttacttttg atactgtttt tggaccagag 300
agtaaacaac ttgatgttta taacttaact gcaagacctt ttattgattc tgtacttgaa 360
ggctacaatg ggactatttt tgcataatgga caaaccggaa caggcaaaac ttttaccatg 420
gaaaggtgtc gagctattcc tgaacttaga ggaataattc cccaatttct ttgctcacia 480
tatttgggcc atatttgc 498

<210> 346
<211> 427
<212> DNA
<213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(427)
 <223> n = A,T,C or G

<400> 346
 agatggcggg cgccgtgaga actttgcagg aacagctgga aaaggccaaa gagagtctta 60
 agaacgtgga tgagaacatt cgcaagctca ccgggcgagg tccgaatgac gtgaggccca 120
 tccaagccag attgctggcc ctttctgggc ctggtggagg tagaggacgt ggtagtttat 180
 tactgaggcg tggattctca gatagtggag gaggaccccc agccaaacag agagaccttg 240
 aaggggcagt cagtaggctg ggcggggagc gtcggaccag aagagaatca cgccaggaaa 300
 gcgaccgga ggatgatgat gttaaaaagc cagcattgca gtcttcant gtagctacct 360
 cccaaagagc gcccacgta gagaccttat ccagggatca aaattttgga tgaaaaagg 420
 gaaagcc 427

<210> 347
 <211> 280
 <212> DNA
 <213> Homo sapien

<400> 347
 cacagaaagt tctccgctcc cagacatggg tccctcggtt tctgcctcg gaagcgcagc 60
 agcaggcatc gtgggaagggt gaagagcttc cctaaggatg acccggtccaa gccgggtccac 120
 ctcacagcct tcttgggata caaggctggc atgactcaca tcgtgcggga agtcgacagg 180
 ccgggatcca aggtgaacaa gaaggagggt gtggaggctg tgaccattgt agagacacca 240
 cccatggttg ttgtgggcat tgtgggctac gtggaaaccc 280

<210> 348
 <211> 411
 <212> DNA
 <213> Homo sapien

<400> 348
 caactatgat gtgcctgaaa aatgggcacg attctatact gcagaagtag ttcttgcatt 60
 ggatgcaatc cattccatgg gttttattca cagagatgtg aagcctgata acatgctgct 120
 ggataaatct ggacatttga agtttagcaga ttttggtagt tgtatgaaga tgaataagga 180
 aggcattggt cagatgtgata cagcgggttg aacacctgat tatatttccc ctgaagtatt 240
 aaaatcccaa ggtggtgatg gttattatgg aagagaatgt gactggtggt cggttgggg 300
 atttttatac gaaatgcttg taggtgatac acctttttat gcagattctt tggttggaac 360
 ttacagtaaa attatgaacc attaaaaatt cacttacctt tcctgatgat a 411

<210> 349
 <211> 408
 <212> DNA
 <213> Homo sapien

<400> 349
 gatgggcacg tctcgggaca actggcacia gcgcccga aaaccggggca agagaaagcc 60
 ctaccacaag aagcgaagt atgagttggg gcgcccagct gccaacacca agattggccc 120
 ccgcccacg cacacagctc gtgtgcgggg aggttaacaag aaataccgtg ccctgagggtt 180
 ggacgtgggg aatttctcct ggggctcaga gtgtgtgact cgtaaaacaa ggatcatcga 240
 tgtgtctac aatgcattca ataacgagct ggttcgtacc aagaccttg tgaagaattg 300
 catcgtgctc atcgacagca caccgtaccg acagtggtag gactccact atgcgctgcc 360
 cctgggcccgc aagaaggagg ccaaactgac ttctgaggaa gaagaaaa 408

<400> 379

tcagatatca	ggtggcttct	tcaaatgatt	tttaagtatc	tcgatgatga	tgaagaacaa	60
agacatcaat	caggattcag	gaagacagct	tttgcgga	atgcttaaag	ggaagcatca	120
aggattgggtg	ttgatatttg	aaagtttaag	agtggatac	ttttattcag	tcaacacatg	180
acaaatgtaa	aaggcactca	tttgttggtc	ctggaagaag	cctggcagca	ttccattcag	240
acatctgccc	tttcatcgtc	ccacttttta	cttattgcag	tcctttcagt	ctgaatattt	300
cctcctgacg	catcttctgc	cgtccgaaat	gactccctgc	tcccagatcc	tgtagccctt	360
attattgaca	ccttttcattt	agaaatttag	cacatgtcac	a		401

<210> 380

<211> 401

<212> DNA

<213> Homo sapien

<400> 380

cctgactctc	tgaggctcat	tttgcagttg	ttgaaattgt	ccccgcagtt	ttcaatcatg	60
tctgaaccaa	tcagagtcct	tgtgactgga	gcagctggtc	aaattgcata	ttcactgctg	120
tacagtattg	gaaatggatc	tgtctttggt	aaagatcagc	ctataattct	tgtgctgttg	180
gatatcacc	ccatgatggg	tgtcctggac	ggtgtcctaa	tggaactgca	agactgtgcc	240
cttcccctcc	tgaaagatgt	catcgcaaca	gataaagaag	acgttgccct	caaagacctg	300
gatgtggcca	ttcttgtggg	ctccatgcca	agaagggaag	gcatggagag	aaaagattta	360
ctgaaagcaa	atgtgaaaat	cttcaaatcc	caggtgtcag	c		401

<210> 381

<211> 401

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(401)

<223> n = A,T,C or G

<400> 381

ggggcttcgc	tggcagtcctg	aacggcaagc	ttgagcaacg	cggtaaaaat	attgcttcgg	60
tgggtgacgc	ggtacagctg	tccaagggcn	ttngtaacgg	gaatgccgaa	gcgtgggaaa	120
aagggagcgg	tggcggaaga	cggggatgag	ctcaggacag	agccagaggc	caagaagagt	180
aagacggccg	caaagaaaaa	tgacaaagag	gcagcaggag	agggcccagc	cctgtatgag	240
gacccccag	atcagaaaac	ctcaccagct	ggcaaacctg	ccacactcaa	gatctgctct	300
tggaatgtgg	atgggcttcg	agcctggatt	aagaagaaag	gattagattg	ggtaaaggaa	360
gaagccccag	atatactgtg	ccttcaagag	accaaattgt	c		401

<210> 382

<211> 491

<212> DNA

<213> Homo sapien

<400> 382

gagcagcccc	cggcgggctga	aagccggggc	agaagtgcctg	gtctcggtcg	ggattccggg	60
cttgggtcca	ccgagggcgg	gactgcggta	ggagggaaga	ggttttgac	gcgctggcct	120
cccgccgctg	tgcattgcag	cattatttca	gttcaaaatg	aactatatgc	ctggcaccgc	180
cagcctcatc	gaggacattg	acaaaaagca	cttggttctg	cttcgagatg	gaaggacact	240
tataggcttt	ttaagaagca	ttgatcaatt	tgcaaactta	gtgctacatc	agactgtgga	300


```

gcgtattcat gtgggcaaaa aatacgggtga tattcctcga gggatttttg tggtcagagg 360
agaaaatgtg gtccctactag gagaaataga cttggaaaag gagagtgaca caccctccca 420
gcaagtatcc attgaagaaa ttctagaaga acaaagggtg gaacagcaga ccaagctgga 480
agcagagaag t 491

```

```

<210> 383
<211> 491
<212> DNA
<213> Homo sapien

```

```

<400> 383
gagtccatct cagcgccctgg aaaatgcagt gaaaaaacct gaagataaaa aggaagtttt 60
cagacccctc aagcctgctg gcgaagtgga tctgaccgca ctggccaaag agcttcgagc 120
agtggaagat gtacggccac ctcacaaagt aacggactac tcctcatcca gtgaggagtc 180
ggggacgacg gatgaggagg acgacgatgt ggagcaggaa ggggctgacg agtccacctc 240
aggaccagag gacaccagag cagcgtcatc tctgaatttg agcaatggtg aaacggaatc 300
tgtgaaaacc atgattgtcc atgatgatgt agaaagttag ccggccatga ccccatccaa 360
ggagggcact ctaatcgctc gccagagtac agttgaccaa aagcgtgcca gccatcatga 420
gagcaatggc tttgccggtc gcattcacct cttgccagat ctcttacagc aaagccattc 480
ctcctccact t 491

```

```

<210> 384
<211> 491
<212> DNA
<213> Homo sapien

```

```

<400> 384
gagcctaate tcaggtggtc cacccgagac cccttgagca ccaaccctag tccccgcgc 60
ggcccccttat tcgctccgac aaggtacaaa aaggctcttg acggcggcgt ggtaggagga 120
cgggagcggg ggcgggaagt tcctgaagg agcgagacag ggagggacag ggcagaggag 180
gagaggaagg cgatgcgacg gacaggcgca cccgctcagg ctgactctcg ggggcgaggt 240
cgagccaggg gcggtgccc tgggggcgag gcgacgctgt ctcaacctcc acctcgcggc 300
ggaacccgag gacaggagcc tcagatgaaa gaaacaatca tgaaccagga aaaactcgcc 360
aaactgcagg cacaagtgcg cattgggtgg aaaggaactg ctgcgagaaa gaagaagggtg 420
gttcatagaa cagccacagc agatgacaaa aaacttcagt tctccttaaa gaagttaggg 480
gtaaacaata t 491

```

```

<210> 385
<211> 483
<212> DNA
<213> Homo sapien

```

```

<400> 385
agccgctgcg aaggaggccg ccgccatgtc tgcgcactctg caatggatgg tcgtgcggaa 60
ctgctccagt ttctgatca agaggaataa gcagacctac agcactgagc ccaataactt 120
gaaggccgcg aattccttcc gctacaacgg actgattcac cgcaagactg tgggcgtgga 180
gcggcgagcc gacggcaaaag gtgtcgtggt ggtcattaag cggagatccg gccagcggaa 240
gcctgccacc tcctatgtgc ggaccaccat caacaagaat gctcgcgcca cgctcagcag 300
catcagacac atgatccgca agaacaagta ccgccccgac ctgcgcagtg cagccatccg 360
cagggccagc gccatcctgc gcagccagaa gcctgtgatg gtgaagagga agcggacccg 420
ccccaccaag agctcctgag cccctgccc ccagagcaat aaagtcagct ggctttctca 480
cct 483

```

```

<210> 386

```

<211> 491
 <212> DNA
 <213> Homo sapien

<400> 386
 aggtggaagg aaaaaacata aatgaagtta atgcacttct tttcctagcc caaaagtcac 60
 tgtgattata tttttttaat gaagtttaga aaaaaagctg ttgtcttctc aattgtaaaa 120
 ttagtttcaa aatgctgctt ctcttatcat tagtctagta attgttgaac ttttctgcaa 180
 actgcattttt acaaaattga aacttggaag ctgtattaac ttttatagtt aaacattgta 240
 ttaataaaac tatactataa taaacagttt ggttttgtat tttttaaatt gtattatcca 300
 gcctttttaa aattaaaagc taaataatga aaataaacca attaaaacat acttttactc 360
 tcagatatac aggtattttac attatgaaaa aactgaacaa agttttaaca atactgagct 420
 ttaagaattt agccagcagg gaaaatttcc aggtttgaga atgttctaata gtaaatattt 480
 aatcataata c 491

<210> 387
 <211> 491
 <212> DNA
 <213> Homo sapien

<400> 387
 ccacaccacc gtgtcccaag tccagccccc tccctccaag gcatcagcac ctgaaccccc 60
 tgcagaagaa gaagtggcaa ctggtacaac ctcagcctct gatgacctgg aagccctggg 120
 tacactgagc ctggggacca cagaggagaa ggcagcagct gaggcggtg tgcccaggac 180
 cattggggcc gagctgatgg agctgggtcg gagaaacact ggctgagcc acgaattatg 240
 ccgggtggcc atcggcatca tagtgggtca catccaggcc tcggtgccg ccagctcacc 300
 agtcatggag caggtcctcc tctcactcgt agagggaag gacctcagca tggccctgcc 360
 ctgagggcag gtctgccacg accagcagag gctggagggtg atctttgcag acctggctcg 420
 ccggaaggac gacgcccagc agcgcagttg ggcactatat gaggatgagg gtgtcatccg 480
 ctgctaccta g 491

<210> 388
 <211> 491
 <212> DNA
 <213> Homo sapien

<400> 388
 gagactatca aactcctgag ccaacaactt aatatgacta gcttacacaa tagcttttat 60
 agtaaagata cctcttttac gactccactt atgactccct aaagcccatg tcgaagcccc 120
 catcgctggg tcaatagtac ttgccgcagt actcttgaaa ctaggcggtg atggtataat 180
 acgacctaca ctcatcttca accccctgac aaaacacata gcctaccctt tcttgtact 240
 atccctatga ggcataatta taacaagctc catctgccta cgacaaacag acctaaaatc 300
 gctcattgca tactcttcaa tcagccacat agccctcgta gtaacagcca ttctcatcca 360
 aacccccctga agcttcaccg gcgcagtcac tctcataatc gccacgggac ttacatcctc 420
 attactattc tgcttagcaa actcaaacta cgaacgcact cacagtcgca tcataatcct 480
 ctctcaagga c 491

<210> 389
 <211> 511
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature

<222> (1)...(511)

<223> n = A,T,C or G

<400> 389

tactgatatc	tctttaatac	tttcatcatt	caagtttgtt	canaacatta	caagaggcat	60
gaaagaaaaa	ataattccat	ttttaaaact	ctgtctgtcc	aaagtataac	atatgaaacc	120
atgccattat	ctnttaggaa	acaaaagcat	tcaaaattaa	tttggtatta	aagttcaaga	180
ttcanactaa	cctcaaagta	cggcattgtc	agtgtttaag	tgcaanaagt	attttcattc	240
caattatttt	acananatgc	tggagtgcg	tgtgcaattt	gaaatattca	aatcctttta	300
ggnttctgaa	ctaagtgttt	aaatgaaaac	tgaaatgctg	catagtttca	gtggctttca	360
atttcctgtt	tgatctcaga	aatatatgga	tgatctttgc	cgtgagctac	ttccatgatt	420
gcaatggcct	tcttcagggc	tttctccctt	gcggtttgtt	gttccaggcc	catgtagagt	480
ctccctagct	tcaaccacat	ggaggccacg	t			511

<210> 390

<211> 1984

<212> DNA

<213> Homo sapien

<400> 390

cctgggggta	gaggctgggg	tgggtggggg	gtaagggggc	agtccttctc	cccttcgacg	60
gcggtccga	gtccagcccc	ttcttcccg	cgctcgctcg	cccggcccc	agccccctca	120
tgagggtgtc	cgtgccgggt	ccggcgcccg	ctgccgcccc	cgcagccggc	cgcgagccct	180
ccacgcccg	cgggggcagc	ggaggcggag	gcgcgctcgc	tgacgcctca	ggcgccggcg	240
tgccgggctc	cgtgcagttg	gcgctgagcg	tcccgacgc	cctgctctac	gccgcgctgt	300
tgccttttgc	ctacctgcag	ctgtggcggc	tgctcctgta	ccgcgagcgg	cggctgagtt	360
accagagcct	ctgcctcttc	ctctgtctcc	tgtgggcagc	gctcaggacc	accctcttct	420
ccgcgcctt	ctgcctcagc	ggctccctgc	ccttgctccg	gccgcccgt	cacctgcact	480
tcttccccc	ctggctgtct	tactgtctcc	cctcctgtct	ccagttctcc	acgtctgtct	540
tcctcaacct	ctacctggcg	gagggtatat	gtaagtcag	atgtgccact	gaacttgaca	600
gacacaaaat	tctactgcat	ttgggcttta	taatggcaag	cctgctcttt	ttagtgggtga	660
acttgacttg	cgcaatgcta	gttcatggag	atgtcccaga	aatcagttg	aagtggactg	720
tgtttggtcg	agcattaatt	aatgatagcc	tgtttattct	ttgtgccatc	tcttttagtgt	780
gttacatatg	caaaattaca	aaaatgtcat	cagctaattg	ctacctcgaa	tcaaagggtga	840
tgtctctgtg	ccagactgtc	atcgtgggct	ctgtagtcat	tcttctgtac	tcttccagag	900
cttggtataa	tttggtgggtg	gtcaccatat	ctcaggatac	attagaaagt	ccatttaatt	960
atggctggga	taatctttca	gataaggctc	atgtagaaga	cataagtga	gaagagtata	1020
tagtatttgg	aatggctctc	tttctgtggg	aacatgtgcc	agcatggtcg	gtgggtactgt	1080
ttttccgggc	acagagatta	aaccagaatt	tggcacctgc	tggcatgata	aatagtcaca	1140
gttatagttc	cagagctttac	tttttcgaca	atccaagacg	atatgatagt	gatgatgacc	1200
tgccaagact	gggaagttca	agagaaggaa	gtttaccaa	ttcgcaaagt	ttgggctgggt	1260
atggcaccat	gactgggtgt	ggcagcagca	gttacacagt	cactccccac	ctgaatggac	1320
ctatgacaga	tactgtctct	ttgtctttta	cttgtagtta	tttagatttg	aacaatcatc	1380
atagcttata	tgtgacacca	caaaactgac	agcatcacca	agtcattgatt	cttgagttgt	1440
ttttcataaa	tgtgtatatt	caatgtgttt	aaattccatc	tacataaaca	ttccattatc	1500
tggtgcaact	gaaaacaaaa	tctggaagtg	tggtgtgtgt	tggtaaataa	cacagctatt	1560
atttttgacc	tcttcatagt	aaaatgaagt	aaaatggaaa	gtttggagta	ggagaaaaga	1620
gagattagat	cttaaggcac	ttgatggcct	ccaaaaatcc	tgactttgga	acatcaaagt	1680
catatgtgca	cttttatctt	tgttctgagt	cactgcagtc	cccaaagtca	tatgccaatg	1740
ttcacactga	aatactgtat	tgtacaccaa	actggaaggc	aattttccta	tgaaaatcaa	1800
agccggtata	ttcattggta	tgctctatac	agatatctta	ataaaaaattt	tatagtgtga	1860
acagtgcaca	gagttaaggc	ataaaaaatgt	atcattcttt	ataaaaaatct	actgaaaatg	1920
tgtaatcatt	gaagacagtt	cttttaagca	tgatttttaa	atagcaactg	aaattcaatc	1980
attt						1984

<400> 391																	
Met	Arg	Val	Ser	Val	Pro	Gly	Pro	Ala	Ala	Ala	Ala	Ala	Ala	Pro	Ala	Ala	
				5					10					15			
Gly	Arg	Glu	Pro	Ser	Thr	Pro	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Gly	Ala	
				20					25					30			
Val	Ala	Ala	Ala	Ser	Gly	Ala	Ala	Val	Pro	Gly	Ser	Val	Gln	Leu	Ala		
				35					40					45			
Leu	Ser	Val	Leu	His	Ala	Leu	Leu	Tyr	Ala	Ala	Leu	Phe	Ala	Phe	Ala		
				50					55					60			
Tyr	Leu	Gln	Leu	Trp	Arg	Leu	Leu	Leu	Tyr	Arg	Glu	Arg	Arg	Leu	Ser		
65					70					75					80		
Tyr	Gln	Ser	Leu	Cys	Leu	Phe	Leu	Cys	Leu	Leu	Trp	Ala	Ala	Leu	Arg		
				85					90					95			
Thr	Thr	Leu	Phe	Ser	Ala	Ala	Phe	Ser	Leu	Ser	Gly	Ser	Leu	Pro	Leu		
				100					105					110			
Leu	Arg	Pro	Pro	Ala	His	Leu	His	Phe	Phe	Pro	His	Trp	Leu	Leu	Tyr		
				115					120					125			
Cys	Phe	Pro	Ser	Cys	Leu	Gln	Phe	Ser	Thr	Leu	Cys	Leu	Leu	Asn	Leu		
				130					135					140			
Tyr	Leu	Ala	Glu	Val	Ile	Cys	Lys	Val	Arg	Cys	Ala	Thr	Glu	Leu	Asp		
145					150					155					160		
Arg	His	Lys	Ile	Leu	Leu	His	Leu	Gly	Phe	Ile	Met	Ala	Ser	Leu	Leu		
				165					170					175			
Phe	Leu	Val	Val	Asn	Leu	Thr	Cys	Ala	Met	Leu	Val	His	Gly	Asp	Val		
				180					185					190			
Pro	Glu	Asn	Gln	Leu	Lys	Trp	Thr	Val	Phe	Val	Arg	Ala	Leu	Ile	Asn		
				195					200					205			
Asp	Ser	Leu	Phe	Ile	Leu	Cys	Ala	Ile	Ser	Leu	Val	Cys	Tyr	Ile	Cys		
				210					215					220			
Lys	Ile	Thr	Lys	Met	Ser	Ser	Ala	Asn	Val	Tyr	Leu	Glu	Ser	Lys	Gly		
225					230					235					240		
Met	Ser	Leu	Cys	Gln	Thr	Val	Ile	Val	Gly	Ser	Val	Val	Ile	Leu	Leu		
				245					250					255			


```

cagccagga gctccccgcc aggccacgcc ccagccaact tgcctcctct cacctctggg 840
aagtgcaaat actcttggtt gacatcccct tcttcagccc tcccagccta ctccccatcc 900
cagcttttag aggatctgct ccaactgtct ctggggcagt tgtgggtcac tgtcccttcc 960
agctgcccc aacaggaagc agagtcacca cgcagcagtg tcccttcttg ggtctgagtt 1020
cctattatag gtagggggccc caccctctgg gcttcccatc agcgacacac acacacttat 1080
ggcaccagcc tggactccag aaaaagggtg tccaggtatt gtgtgtatgc atttagttgt 1140
gcacacacaa atatgctcct atactggcat taggcgtctc ctcatccctc accctgacct 1200
ttctcctgtc cttttcttgg ctggaagaag ttggcctcct gggagtgtag ttttctgttt 1260
taaatcccc acccttggtt gggctcagtg gctcaccctt gtaatcccag cactttggga 1320
ggccaaggcg ggtcgattac ttgaggtcag gagttcaaga ccagcctggc caacattgtg 1380
aaaccccatc tctgccaaaa atacaaaagt tagccgggcg tagtggcaca tgctgtaat 1440
cccagctacc cggggagggt gaggcaggag aattgcttga actcagaagg cggaggctgc 1500
agtgagccga gatcgtgcca ctgcactcca gcctgggtcaa cagagcaaga ctccatctcg 1560
aaaaaaaaa aaaaaaaact cgag 1584

```

```

<210> 393
<211> 191
<212> PRT
<213> Homo sapiens

```

```

<400> 393
Met Gly Lys Ser Cys Lys Val Val Val Cys Gly Gln Ala Ser Val Gly
      5                      10                      15

Lys Thr Ser Ile Leu Glu Gln Leu Leu Tyr Gly Asn His Val Val Gly
      20                      25                      30

Ser Glu Met Ile Glu Thr Gln Glu Asp Ile Tyr Val Gly Ser Ile Glu
      35                      40                      45

Thr Asp Arg Gly Val Arg Glu Gln Val Arg Phe Tyr Asp Thr Arg Gly
      50                      55                      60

Leu Arg Asp Gly Ala Glu Leu Pro Arg His Cys Phe Ser Cys Thr Asp
      65                      70                      75                      80

Gly Tyr Val Leu Val Tyr Ser Thr Asp Ser Arg Glu Ser Phe Gln Arg
      85                      90                      95

Val Glu Leu Leu Lys Lys Glu Ile Asp Lys Ser Lys Asp Lys Lys Glu
      100                     105                     110

Val Thr Ile Val Val Leu Gly Asn Lys Cys Asp Leu Gln Glu Gln Arg
      115                     120                     125

Arg Val Asp Pro Asp Val Ala Gln His Trp Ala Lys Ser Glu Lys Val
      130                     135                     140

Lys Leu Trp Glu Val Ser Val Ala Asp Arg Arg Ser Leu Leu Glu Pro
      145                     150                     155                     160

Phe Val Tyr Leu Ala Ser Lys Met Thr Gln Pro Gln Ser Lys Ser Ala
      165                     170                     175

```

"000000" 000000

Phe Pro Leu Ser Arg Lys Asn Lys Gly Ser Gly Ser Leu Asp Gly
 180 185 190

<210> 394
 <211> 1937
 <212> DNA
 <213> Homo sapiens

<400> 394
 ccggttcccc cagctctggg taccgggctc tgcctgcggt cgccatgatg ggccatcgctc 60
 cagtgtctcgt gctcagccag aacacaaagc gtgaatccgg aagaaaagtt caatctggaa 120
 acatcaatgc tgccaagact attgcagata tcatccgaac atgtttggga cccaagtcca 180
 tgatgaagat gcttttggac ccaatgggag gcattgtgat gaccaatgat ggcaatgcc 240
 ttcttcgaga gattcaagtc cagcatccag cggccaagtc catgatcgaa attagccgga 300
 cccaggatga agagggttga gatgggacca catcagtaat tattcttgca ggggaaatgc 360
 tgtctgtagc tgagcacttc ctggagcagc agatgcaccc aacagtgggtg atcagtgcctt 420
 accgcaaggc attggatgat atgatcagca ccctaaagaa aataagtatc ccagtcgaca 480
 tcagtgcacg tgatatgatg ctgaacatca tcaacagctc tattactacc aaagccatca 540
 gtccgtgggtc atctttgggt tgcaacattg ccctggatgc tgtcaagatg gtacagtttg 600
 aggagaatgg tcggaaagag attgacataa aaaaatatgc aagagtggaa aagataacctg 660
 gaggcacatc tgaagactcc tgtgtcttgc gtggagtcac gattaacaag gatgtgacct 720
 atccacgtat gcggcgctat atcaagaacc ctgcgattgt gctgctggat tcttctctgg 780
 aatacaagaa aggagaaagc cagactgaca ttgagattac acgagaggag gacttcaccc 840
 gaattctcca gatggaggaa gagtacatcc agcagctctg tgaggacatt atccaactga 900
 agcccgatgt ggtcatcact gaaaaggcca tctcagattt agctcagcac taccttatgc 960
 gggccaatat cacagccatc cgcagagtcc ggaagacaga caataatcgc attgctagag 1020
 cctgtggggc ccggatagtc agccgaccag aggaactgag agaagatgat gttggaacag 1080
 gagcaggcct gttggaaatc aagaaaattg gagatgaata ctttactttc atcactgact 1140
 gcaaagaccc caaggcctgc accattctcc tccggggggc tagcaaagag attctctcgg 1200
 aagtagaacg caacctccag gatgccatgc aagtgtgtcg caatgttctc ctggaccctc 1260
 agctgggtgc aggggttggg gcctccgaga tggctgtggc ccatgccttg acagaaaaat 1320
 ccaaggccat gactggtgtg gaacaatggc catacagggc tgttgcccag gccctagagg 1380
 tcattcctcg taccctgatc cagaactgtg gggccagcac catccgtcta cttacctccc 1440
 ttccggccaa gcacacccag gagaactgtg agacctgggg tgtaaattgt gagacgggta 1500
 ctttgggtga catgaaggaa ctgggcatat gggagccatt ggctgtgaag ctgcagactt 1560
 ataagacagc agtggagacg gcagttctgc tactgcgaat tgatgacatc gtttcaggcc 1620
 acaaaaagaa aggcgatgac cagagccggc aaggcggggc tcctgatgct ggccaggagt 1680
 gagtgttagg caaggctact tcaatgcaca gaaccagcag agtctccctc tttcctgagc 1740
 cagagtgcc aagaactgt ggacgtcttt gttcagaagg gatcagggtg gggggcagcc 1800
 cccagtcctt ttctgtccca gctcagtttt ccaaaagaca ctgacatgta attcttctct 1860
 attgtaaggt ttccatttag tttgcttccg atgattaaat ctaagtcatt tgaaaaaaaa 1920
 aaaaaaaaaa actcgag 1937

<210> 395
 <211> 1675
 <212> DNA
 <213> Homo sapiens

<400> 395
 gcgcgaatcg cggctcgcgag ccattggagga ggaggcatcg tccccggggc tgggctgcag 60
 caagccgcac ctggagaagc tgaccctggg catcacgcgc atcctagaat cttccccagg 120
 tgtgactgag gtgaccatca tagaaaagcc tcctgctgaa cgtcatatga tttcttctctg 180
 ggaacaaaag aataactgtg tgatgcctga agatgtgaag aacttttacc tgatgaccaa 240
 tggcttccac atgacatgga gtgtgaagct ggatgagcac atcattccac tgggaagcat 300

0956710"095600

ggcaattaac agcatctcaa aactgactca gctcaccag tcttccatgt attcacttcc 360
 taatgcaccc actctggcag acctggagga cgatacacat gaagccagtg atgatcagcc 420
 agagaagcct cactttgact ctgcagtgat gatatttgag ctggattcat gcaatggcag 480
 tgggaaagt tgccttgtct acaaaagtgg gaaaccagca ttagcagaag aactgagat 540
 ctggttcctg gacagagcgt tatactggca ttttctcaca gacaccttta ctgcctatta 600
 ccgctgctc atcaccacc tgggcctgcc ccagtggcaa tatgccttca ccagctatgg 660
 cattagccca caggccaagc aatgggttcag catgtataaa cctatcacct acaacacaaa 720
 cctgctcaca gaagagaccg actcctttgt gaataagcta gatcccagca aagtgtttaa 780
 gagcaagaac aagatcgtaa tccccaaaaa gaaagggcct gtgcagcctg caggtggcca 840
 gaaagggccc tcaggaccct ccgggtccctc cacttctctc acttctaaat cctcctctgg 900
 ctctggaaac cccaccggga agtgagcacc cctccctcca actccctacc agctccagag 960
 tgggtggttc catgcacaga tggccctagg ggtgacctcc agttttgcgt gtggaccgta 1020
 ggctctttc tagttgaatg accaaaattg taaggctttt agtcccaccg acattagcca 1080
 ggctcgtagt gaggcctcca gagcagggtg tgcctctgga agcaatgggg 1140
 aatttggaat ctgtgtgaag tgcccaaata agtctgagt ctttctctt cttcaacact 1200
 caaccctcaa tcccttagca ctgattgatt agagaggtcc cccaaagaaa ccactgggtt 1260
 tgacccatga agcattagaa ctgcattgtt cattcaggag ccactagtca catatgacta 1320
 tttaaattta aagtaaattg tatgaaaaat tcatttcttc aattgcatta gccacatttt 1380
 gagtattcat gtggctggta gattctgtat tagcaciaag atatggaaca tttccatcac 1440
 cacagaaagt tctgttggac agcactgcat tagaatatt tcatactgct cttcctcaat 1500
 taatttttgt tgtaaatgtt gatgtcttca ttggatgggt cataatgttc catgaaacct 1560
 ctcaagtaca caattgtatg ttctttgtat cccttaccac aaatatctcg ctctgctcat 1620
 ttcttttgca gcttctata aagtttgtct tcctcatcaa aaaaaaaaaa aaaaa 1675

<210> 396

<211> 559

<212> PRT

<213> Homo sapiens

<400> 396

Gly Ser Pro Ser Ser Gly Tyr Pro Ala Leu His Arg Val Ala Met Met
 5 10 15

Gly His Arg Pro Val Leu Val Leu Ser Gln Asn Thr Lys Arg Glu Ser
 20 25 30

Gly Arg Lys Val Gln Ser Gly Asn Ile Asn Ala Ala Lys Thr Ile Ala
 35 40 45

Asp Ile Ile Arg Thr Cys Leu Gly Pro Lys Ser Met Met Lys Met Leu
 50 55 60

Leu Asp Pro Met Gly Gly Ile Val Met Thr Asn Asp Gly Asn Ala Ile
 65 70 75 80

Leu Arg Glu Ile Gln Val Gln His Pro Ala Ala Lys Ser Met Ile Glu
 85 90 95

Ile Ser Arg Thr Gln Asp Glu Glu Val Gly Asp Gly Thr Thr Ser Val
 100 105 110

Ile Ile Leu Ala Gly Glu Met Leu Ser Val Ala Glu His Phe Leu Glu
 115 120 125

Gln	Gln	Met	His	Pro	Thr	Val	Val	Ile	Ser	Ala	Tyr	Arg	Lys	Ala	Leu
130						135			140						
Asp	Asp	Met	Ile	Ser	Thr	Leu	Lys	Lys	Ile	Ser	Ile	Pro	Val	Asp	Ile
145			150						155			160			
Ser	Asp	Ser	Asp	Met	Met	Leu	Asn	Ile	Ile	Asn	Ser	Ser	Ile	Thr	Thr
			165						170			175			
Lys	Ala	Ile	Ser	Arg	Trp	Ser	Ser	Leu	Ala	Cys	Asn	Ile	Ala	Leu	Asp
			180			185						190			
Ala	Val	Lys	Met	Val	Gln	Phe	Glu	Glu	Asn	Gly	Arg	Lys	Glu	Ile	Asp
195						200						205			
Ile	Lys	Lys	Tyr	Ala	Arg	Val	Glu	Lys	Ile	Pro	Gly	Gly	Ile	Ile	Glu
210						215			220						
Asp	Ser	Cys	Val	Leu	Arg	Gly	Val	Met	Ile	Asn	Lys	Asp	Val	Thr	His
225			230						235			240			
Pro	Arg	Met	Arg	Arg	Tyr	Ile	Lys	Asn	Pro	Arg	Ile	Val	Leu	Leu	Asp
			245						250			255			
Ser	Ser	Leu	Glu	Tyr	Lys	Lys	Gly	Glu	Ser	Gln	Thr	Asp	Ile	Glu	Ile
			260			265						270			
Thr	Arg	Glu	Glu	Asp	Phe	Thr	Arg	Ile	Leu	Gln	Met	Glu	Glu	Glu	Tyr
275						280						285			
Ile	Gln	Gln	Leu	Cys	Glu	Asp	Ile	Ile	Gln	Leu	Lys	Pro	Asp	Val	Val
290						295			300						
Ile	Thr	Glu	Lys	Gly	Ile	Ser	Asp	Leu	Ala	Gln	His	Tyr	Leu	Met	Arg
305			310						315			320			
Ala	Asn	Ile	Thr	Ala	Ile	Arg	Arg	Val	Arg	Lys	Thr	Asp	Asn	Asn	Arg
			325			330						335			
Ile	Ala	Arg	Ala	Cys	Gly	Ala	Arg	Ile	Val	Ser	Arg	Pro	Glu	Glu	Leu
			340			345						350			
Arg	Glu	Asp	Asp	Val	Gly	Thr	Gly	Ala	Gly	Leu	Leu	Glu	Ile	Lys	Lys
355						360						365			
Ile	Gly	Asp	Glu	Tyr	Phe	Thr	Phe	Ile	Thr	Asp	Cys	Lys	Asp	Pro	Lys
370						375			380						
Ala	Cys	Thr	Ile	Leu	Leu	Arg	Gly	Ala	Ser	Lys	Glu	Ile	Leu	Ser	Glu
385			390						395			400			
Val	Glu	Arg	Asn	Leu	Gln	Asp	Ala	Met	Gln	Val	Cys	Arg	Asn	Val	Leu
			405						410			415			

Thr Arg Lys
305

```
<210> 398
<211> 416
<212> DNA
<213> Homo sapiens
```

```
<400> 398
agaattcggc acgaggattg cctatctcca gtgcaacaac catcaagtgt gctgaaagtc 60
ttcagccggt tgctgcagca gtggaagaaa gggctacagg tccagtcttg ataagcaccg 120
ccgactttga ggggcctatg cccagtgcgc cccagaagc tgaaagtctt cttgcctcaa 180
ccagcaagga ggagaaggat gaatgtgctc tcatttccac tagcatagca gaagaatgtg 240
aggtctctgt ttccggtgta gttgttgaaa gtgaaaatga gcgagctggc acagtcatgg 300
```

```

<400> 401
ggnacgagga atcatggcgg ctgcgctggt cgtgctgctg ggattcgcgc tgctgggcac 60
ccacggagcc tccggggctg ccggcacagt cttcactacc gtagaagacc ttggctccaa 120
gatactcttc acctgtctct tgaatgacag cgccacagag gtcacagggc accgctggct 180
gaaggggggc gtggtgctga aggaggacgc gctgccgggc cagaaaacgg agttcaaggt 240
ggactccgac gaccagtggg gagagtactc ctgcgtcttc ctcccogagc ccatgggcac 300
ggccaacatc cagctccacg ggctcccag agtgaaggcc gtgaagtgtg cagaacacat 360
caacgagggg gagacggcca tgctggtctg caagtacagag tccgtgccac ctgtcactga 420
ctgggectgg tac                                     483

```

<210> 402
 <211> 434
 <212> DNA
 <213> Homo sapiens

<400> 402
 ggcacgaggc tcggactgag caggactttc cttatcccag ttgattgtgc agaatacact 60
 gcctgtcgct tgtcttctat tcacccatggc ttcttctgat atccagggtga aagaactgga 120
 gaagcgtgcc tcaggccagg cttttgagct gattctcagc cctcgggtcaa aaggatctgt 180
 tccagaattc cccctttccc ctccaaagaa gaaggatctt tccctggagg aaattcagaa 240
 gaaattagaa gctgcagaag aaagacgcaa gtcccatgaa gctgaggtct tgaagcagct 300
 ggctgagaaa cgagagcacg agaaagaagt gcttcagaag gcaatagaag agaacaacaa 360
 cttcagtaaa atggcagaag agaaactgac ccacaaaatg gaagctaata aagagaaccg 420
 agaggcacia atgg 434

<210> 403
 <211> 435
 <212> DNA
 <213> Homo sapiens

<400> 403
 ggcacgagga actgctgttg ccattcaaac cattgaggag catcctgcat cttttgactg 60
 gagctctttt aagccaatgg gatttgaagt atcatttctg aagtttcttg aggagtctgc 120
 agtgaagcag aagaaaaata ctgacaaaga ccacccgaat actggaaaca aaaaaggatc 180
 ccattcaaat tcaagaaaaa atattgataa gactgctgtg actagtggaa atcatgtatg 240
 tccttgtaaa gaaagcgaaa cgtttgtaca gtttgccaat ccacacagc ttcagtgcag 300
 tgataatgta aaaattgttt tagacaagaa tcttaaagat tgcactgagc ttgtcttaaa 360
 gcaacttcag gaaatgaaac ctaccgtcag tctgaaaaaa cttgaagtac attcaaata 420
 tccagatatg tctgt 435

<210> 404
 <211> 416
 <212> DNA
 <213> Homo sapiens

<400> 404
 aaagaattcg gcacgaggeg cgctccgcc acgaccaccg ccgcctcctg ccctgcagcc 60
 accgccaccg cctgtgtcgc cgccgcctcg ggaccggctg tatgattagg ccacaatctt 120
 caatgagtaa acatattcct caattctgtg gtgttcttgg tcacacattt atggagtctt 180
 tgaagggcag tggagattac tgccaggcac agcacgacct ctatgcagac aagtgaactg 240
 tagaaactga ttactgctcc accaagaagc ccccataaga gtggttatcc tggacacaga 300
 agtggtgaat tgaaatccac agagcatttt acaagagttc tgacctggat ggggtaaac 360
 tcagtgcact tcttttctgt tggcctcagt attactggat tgaagaattg ctgctt 416

<210> 405
 <211> 435
 <212> DNA
 <213> Homo sapiens

<400> 405
 ggcacgaggg ctgccggagg gtcgttttaa agggcccgcg cgttgccgcc ccctcggccc 60
 gccatgctgc tatccgtgcc gctgctgctc ggctcctcgc gcctggccgt cgccgagcct 120
 gccgtctact tcaaggagca gtttctggac ggagacgggt ggacttcccg ctggatcgaa 180
 tccaaacaca agtcagattt tggcaaattc gttctcagtt ccggcaagtt ctacggtgac 240

000250 "02" 09960

```

gaggagaaag ataaaggttt gcagacaagc caggatgcac gcttttatgc tctgtcggcc 300
agtttcgagc ctttcagcaa caaaggccag acgctgggtg tgcagttcac ggtgaaacat 360
gagcagaaca tgcactgtgg gggcggtat gtgaagctgt ttctaatag tttggaccag 420
acagacatgc acgga 435

```

```

<210> 406
<211> 424
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(424)
<223> n = A,T,C or G

```

```

<400> 406
gcccaaacc actccacctt actaccagac aaccttagcc aaaccattta cccaaataaa 60
gtataggcga tagaaattga aacctggcgc aatagatata gtaccgcaag ggaaagatga 120
aaaattataa ccaagcataa tatagcaagg actaaccctt ataccttctg cataatgaat 180
taactagaaa taactttgca aggagagcca aagctaagac ccccgaaacc agacgagcta 240
cctaagaaca gctaaaagag cacacccgtc tatgtagcaa aatagtggga agatttatag 300
gtagaggcga caaacctacc gagcctggtg atagctggtt gtccaagata gaatcttagt 360
tcaactttaa atttgccac agaaccctct aaatcccctt gnaaatttaa ctgntagtc 420
aaag 424

```

```

<210> 407
<211> 423
<212> DNA
<213> Homo sapiens

```

```

<400> 407
gtcctaccg gcgcacgtgg tgccgcgct gctgcctccc gctcgccctg aaccagtg 60
ctgcagccat ggctcccggc cagctcgctt tatttagtgt ctctgacaaa accggccttg 120
tggaatttgc aagaaacctg accgctcttg gtttgaatct ggctcgcttc ggagggactg 180
caaaagctct cagggatgct ggtctggcag tcagagatgt ctctgagttg acgggatttc 240
ctgaaatgtt ggggggacgt gtgaaaactt tgcacctcgc agtccatgct ggaatcctag 300
ctcgtaatat tccagaagat aatgctgaca tggccagact tgatttcaat cttataagag 360
ttgttgctg caatctctat ccctttgtaa agacagtggc ttctccaggt gtaagtgttg 420
agg 423

```

```

<210> 408
<211> 424
<212> DNA
<213> Homo sapiens

```

```

<400> 408
gaaaaaaaa atgttactga attctataag atgtgtggga atctcaccta tcaaaaaatag 60
gtaaaaagag cctccaaacc tgctttgatt ttattcacct attcttttag gccaggaact 120
aatttacctc tcaactatct gttccctctt gctatcttgt ggagtctcta aagacaaagg 180
tataaagagc ttttggtagg tgaattaata atcaactaga tggcatttcc aaatgggatt 240
gcacatactg tggggcaagt cccaagtga cttcaaagtg agacgtttat ttgagtaatc 300
cttcagatt aacaataatc ataatagcag ttaccacttc ctgagtactt tctatatgcc 360
atgtattgag cttgctcact tctttatgtg gattcttatt taatcttaat accaagatga 420
ggtg 424

```

<210> 409
 <211> 398
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(398)
 <223> n = A,T,C or G

<400> 409
 gctcgactct tagcttgtcg gggacggtaa ccgggacccg gtgtctgctc ctgtcgccctt 60
 cgctcctaa tccctagcca ctatgcgtga gtgcatctcc atccacgttg gccaggctgg 120
 tgtccagatt ggcaatgcct gctgggagct ctactgcctg gaacacggca tccagcccga 180
 tggccagatg ccaagtgaca agaccattgg gggaggagat gactccttca acaccttctt 240
 cagtgaagcg ggcgctggca agcacgtgcc ccgggctgng tttgtagact tggaaaccac 300
 agtnattgat gaagntcgna ctggcaccta cccgcaggtc ttncaccctg ancanntcat 360
 nacaggcaag gaagatgctg ncaaataact atgcccgga 398

<210> 410
 <211> 423
 <212> DNA
 <213> Homo sapiens

<400> 410
 gccccacccc acctgcccgc tgcggctctc cgcgggagat ctcaccgttc tggagacagg 60
 gctcgctcgc tctcacgctg cccggccagc ccgcttctct gcccgagacc atgaatctca 120
 gtagcgccag tagcacggag gaaaaggcag tgacgaccgt gctctggggc tgcgagctca 180
 gtcaggagag gcggacttgg accttcagac cccagctgga ggggaagcag agctgcaggc 240
 tgttgcttca tacgatttgc ttggggggaga aagccaaaga ggagatgcat cgcgtggaga 300
 tcctgcccc agcaaaccag gaggacaaga agatgcagcc ggtcaccatt gcctcactcc 360
 aggcctcagt cctccccatg gtctccatgg taggagtgca gctttctccc ccagttactt 420
 tcc 423

<210> 411
 <211> 424
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(424)
 <223> n = A,T,C or G

<400> 411
 gcggaggcga ctagcggcgg cgggagcggc gccgagaggc cgtgcgggac gcgggcgcca 60
 ggaccggccg aacgcagagg ttgattcttc accacactga aaccattagg aaaaatcctt 120
 gtgggttaaca gcagaggctt cagagtgtaa cctgtactcg ggcctagaaa ttatttaaaa 180
 tggcgactga tacgtctcaa ggtgaactcg tccatcctaa ggcactccca cttatagtag 240
 gagctcagct gatccacgcg gacaagttag gtgagaaggt agaagatagc accatgccga 300
 ttcgtcgaac tgtgaattct acccgggaaa ctctcctccaa aagcaagctt gctgaagggg 360
 aggaagaaan gccagaacca gacataagtt cagaggaatc tgtctccact gtagaagaac 420
 aaga 424

<210> 412
 <211> 430
 <212> DNA
 <213> Homo sapiens

<400> 412
 ggcacgaggg gaagccggcg ccagttcgcg gggctccggg ccgccactca gagctatgag 60
 ctacggccgc cccctcccgc atgtggaggg tatgacctcc ctcaagggtg acaacctgac 120
 ctaccgcacc tcgcccgcgc cgctgagggc cgtcttcgag aagtacgggc gcgtcggcga 180
 cgtgtacatc ccgcgggata gctacaccaa ggagtccgcg ggcttcgcct tcgttcgctt 240
 tcacgacaag cgcgacgctg aggacgctat ggatgccatg gacggggccg tgctggacgg 300
 ccgcgagctg cgggtgcaaa tggcgcgcta cggccgcccc ccggactcac accacagccg 360
 ccggggaccg ccaccccgcg ggtacggggg cggtggttac ggacgccgga gccgcagccc 420
 taggcggcgt 430

<210> 413
 <211> 429
 <212> DNA
 <213> Homo sapiens

<400> 413
 ggcacgaggt cggcccggcc atcttgtggg aagagctgaa gcaggcgctc ttggctcggc 60
 gcggcccgt gcaatccgtg gaggaacgcg ccgccgagcc accatcatgc ctgggcactt 120
 acaggaaggc ttccgctgcg tggtcaccaa ccgattcgac cagttatttg acgacgaatc 180
 ggaccccttc gaggtgctga aggcagcaga gaacaagaaa aaagaagccg gcggggggcg 240
 cgttgggggc cctggggcca agagcgcagc tcaggccgcg gccagacca actccaacgc 300
 ggcaggcaaa cagctgcgca aggagtcca gaaagaccgc aagaacccgc tgccccccag 360
 cgttgccgtg gttgacaaga aagaggagac gcagccgccc gtggcgctta agaaagaagg 420
 aataagacg 429

<210> 414
 <211> 429
 <212> DNA
 <213> Homo sapiens

<400> 414
 ggcacgagga cgggcccggc tgccggcccc cgctctgccc tgcataataa aatggctaata 60
 caggtgaatg gtaatgcggt acagttaaaa gaagaggaag aaccaatgga tacttccagt 120
 gtaactcaca cagaacacta caagacactg atagaggcag gcctcccaca gaagggtggc 180
 gaaagacttg atgaaatatt tcagacagga ttggtagctt atgtcgatct tgatgaaaga 240
 gcaattgatg ctctcaggga atttaataaa gaaggagctc tgtctgtact acagcagttc 300
 aaggaaaagt acttatcaca tgttcagaac aaaagtgcac ttttatgttg agttatgaag 360
 acctacaggc agagagagaa acaggggagc aagggtgcaag agtccacaaa gggacctgat 420
 gaagcgaag 429

<210> 415
 <211> 398
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(398)

000250"0229960

<223> n = A,T,C or G

<400> 415

```
gcggtcgtaa gggctgagga tttttggtcc gcacgctcct gtcctgact caccgctgtt 60
cgctctcgcc gaggaacaag tcggtcagga agcccgcgcg caacagccat ggcttttaag 120
gataccggaa aaacacccgt ggagccggag gtggcaattc accgaattcg aatcaccta 180
acaagccgca acgtaaaatc cttggaaaag gtgtgtgctg acttgataag aggcgcaaaa 240
gaaaagaatc tcaaagtact ttgagaatca ctacaagaaa aactccttgt ggtgaagggt 300
ctaagacgtg ggatcgtttc cagatgagaa ttcacaagcg actcattgac ttgcacagtc 360
cttctgagat tgtaagcan attacttcca tcantatt 398
```

<210> 416

<211> 269

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(269)

<223> n = A,T,C or G

<400> 416

```
gccgaggcag gaagctgtga gtgcgcgggt gcggggtcgc attgtggcta cggctttgcg 60
tccccggcgg gcagccccag gctgggtccc gcctccgctc tccccaccgg cggggaaagc 120
agctggtgtg ggaggaaagg ctccatcccc cgccccctct ctcccgtgtg ttgctggcan 180
gatcttttgg cagtcctgtg gnetcnctcc ccgnccggat cctnctgacc ctganattcn 240
nggtntnacn nnccgtncac gccttgntt 269
```

<210> 417

<211> 408

<212> DNA

<213> Homo sapiens

<400> 417

```
ggccgggaga accgttcgcg gaggaaaggc gaactagtgt tgggatggcc accaaactggg 60
ggagcctctt gcaggataaa cagcagctag aggagctggc acggcaggcc gtggaccggg 120
ccctggctga gggagtattg ctgaggacct cacaggagcc cacttctctg gaggtggtga 180
gctatgcccc attcacgtct tccccctcac tgggtccccag tgccctgctg gagcaagcct 240
atgctgtgca gatggacttc aacctgctag tggatgctgt cagccagaac gctgccttcc 300
tggagcaaac tctttccagc accatcaaac aggatgactt taccgctcgt ctctttgaca 360
tcacaaagca agtccataaa gagggcattg cccagactgt gttcctgg 408
```

<210> 418

<211> 402

<212> DNA

<213> Homo sapiens

<400> 418

```
gagccgggca gccgcttccc gcccccgagc aggagccggt gcgagcggag cagagccgag 60
gtcggggcgc gagcggagcc ggctgagcgg gcgccgagct cccgccatgg cccggaacac 120
gctgtcctcg cgttcccgcc ggggtggacat cgacgaattt gacgagaaca aatttggtga 180
cgagcaggag gaggcggcgg cggcggcggc ggagccaggc ccggacccga gcgaggtgga 240
cgggctcctg cggcaagggg acatgcttcg ggcattccat gcagccttgc ggaactctcc 300
cgtcaacacc aagaatcaag ctgtgaagga gcgagcccag ggcgtggtgc tgaaagtgct 360
```

cacaaacttc aagagcagtg agattgagca ggctgtgcag tc

402

<210> 419

<211> 406

<212> DNA

<213> Homo sapiens

<400> 419

```

gcccgggccta gcggcctggg ttgggctttg tagctgctcc gcaggcccag cccggggcgc 60
gctcgagag tcctaggcgg tgcgcggcct cctgcctcct ccctcctcgg cggtcgcggc 120
ccgcccgcct ccgcgggtgcc tgccttcgct ctcagggtga ggagctcaag cttgggaaaa 180
tggtgtgcat tccttgatc gtcattccag ttctgctctg gatctacaaa aaattcctgg 240
agccatatat atacctctg gtttccccct tcgttaagtc gtatatggcc taaaaaaga 300
attcaaagaa atccaatgat ccaaacaaaa gggcaaaagt aaaaactttt aaaggggtgc 360
aagaacattg aaatgggaat tacccaacca aaaaaggga cccaac 406

```

<210> 420

<211> 371

<212> DNA

<213> Homo sapiens

<400> 420

```

cagccatcgt ggtgtgttct tgactccgct gctcgccatg tcttctcaca agactttcag 60
gattaagcga ttcttgcca agaaacaaaa gcaaaatcgt cccattcccc agtggattcg 120
gatgaaaact ggaaataaaa tcaggtacaa ctccaaaagg agacattgga gaagaaccaa 180
gctgggtcta taaggaattg cacatgagat ggcacacata tttatgctgt ctgaagggtca 240
cgatcatgtt accatatcaa gctgaaaatg tcaccactat ctggagattt cgacgtgttt 300
tcctctctga atctgttatg aacacgttgg ttggctggat tcagtaataa atatgtaagg 360
cctttctttt t 371

```

<210> 421

<211> 51

<212> PRT

<213> Homo sapiens

<400> 421

```

Met Ser Ser His Lys Thr Phe Arg Ile Lys Arg Phe Leu Ala Lys Lys
                5                10                15

```

```

Gln Lys Gln Asn Arg Pro Ile Pro Gln Trp Ile Arg Met Lys Thr Gly
                20                25                30

```

```

Asn Lys Ile Arg Tyr Asn Ser Lys Arg Arg His Trp Arg Arg Thr Lys
                35                40                45

```

```

Leu Gly Leu
                50

```

<210> 422

<211> 12308

<212> DNA

096610-09000

$\langle 400 \rangle$ 422

ttgtttttcta	gcagtgacaa	gttcacttttg	aatcaggata	tgtgtgtagt	ttgtggcagt	60
tttgccaag	gagcagaag	aagattactt	gcctgttctc	agtgtggtca	gtgttaccat	120
ccatactgtg	tcagtattaa	gatcactaaa	gtgggttctta	gcaaaggttg	gaggtgtctt	180
gagtgcactg	tgtgtgaggc	ctgtgggaag	gcaactgacc	caggaagact	cctgctgtgt	240
gatgactgtg	acataagtta	tcacacctac	tgcctagacc	ctccattgca	gacagttccc	300
aaaggaggct	ggaagtgcaa	atggtgtggt	tggtgcagac	actgtggagc	aacatctgca	360
ggtctaagat	gtgaatggca	gaacaattac	acacagtgcg	ctccttgtgc	aagcttatct	420
tcctgtccag	tctgtatcgc	aaactataga	gaagaagatc	ttattctgca	atgtagacaa	480
tgtgatagat	ggatgcatgc	agtttgtcag	aacttaaata	ctgaggaaga	agtggaaat	540
gtagcacaca	ttggttttga	ttgtagcatg	tgcagacctt	atatgctctgc	gtctaattgtg	600
ccttcctcag	actgtgtgta	atcttcactt	gtagcacaaa	ttgtcacaaa	agtaaaagag	660
ctagaccac	ccaagactta	taccaggat	ggtgtgtggt	tgactgaatc	agggatgact	720
cagttacaga	gcctcacagt	tacagttcca	agaagaaaaac	ggtcaaaacc	aaaattgaaa	780
ttgaagatta	taaatacagaa	tagcgtggcc	gtccttcaga	ccctccaga	catccaatca	840
gagcattcaa	gggatggtga	aatggatgat	agtcgagaag	gagaacttat	ggatttgtgat	900
ggaaaatcag	aatctagtcc	tgagcgggaa	gctgtggatg	atgaaactaa	gggagtggaa	960
ggaacagatg	gtgtcaaaaa	gagaaaaagg	aaaccataca	gaccaggtat	tggtggattt	1020
atggtgoggc	aaagaagtcg	aactgggcaa	gggaaaacca	aaagatctgt	gatcagaaaa	1080
gattcctcag	gctctatttc	cgagcagtta	ccttgccagag	atgatggctg	gagtgagcag	1140
ttaccagata	ctttagttga	tgaatctggt	tctgttactg	aaagcactga	aaaaataaag	1200
aagagatacc	gaaaaaggaa	aaataagctt	gaagaaaactt	tccttgcta	tttacaagaa	1260
gctttctttg	gaaaagatct	tctagatata	agtagacaaa	gcaacataag	tttagtaaat	1320
ctgtcagaag	atggagctca	gcttttatat	aaacacaaa	tgaacacagg	tttctgtgat	1380
ccttccttag	atccactact	tagtctatcc	tcgggtccaa	caaaatctgg	aactcacggt	1440
cctgctgatg	accattagc	tgatatttct	gaagttttaa	acacagatga	tgacattctt	1500
ggaataattt	cagatgatct	agcaaaatca	gttgatcatt	cagatattgg	tcctgtcact	1560
gatgatecct	cctctttgcc	tcagccaaat	gtcaatcaga	gttcacgacc	attaagttaa	1620
gaacagctag	atgggatcct	cagtcctgaa	ctagacaaaa	tggtcacaga	tgagcaatt	1680
cttgaaaaat	tatataaaat	tcagagcctt	ggcggaagg	atgttgaaga	cttattttaca	1740
gctgtactta	gtcctgcgaa	cactcagcca	actccattgc	cacagcctcc	cccaccaaca	1800
cagctgttgc	caatacacaa	tcaggatgct	ttttcacgga	tgccctctcat	gaatggcctt	1860
attggatcca	gtcctcatct	cccacataat	tctttgccac	ctggaagcgg	actgggaact	1920
ttctctgcaa	ttgcacaatc	ctcttatcct	gatgccaggg	ataaaaattc	agcctttaat	1980
ccaatggcaa	gtgatecctaa	caactcttgg	acatcatcag	ctcccactgt	ggaaggagaa	2040
aatgacacaa	tgtcgaatgc	ccagagaagc	acgcttaagt	gggagaaaga	ggaggtctgt	2100
ggtgaaattg	caactgttgc	ccaggtcttc	tacaccaata	ttaaattccc	caacttaag	2160
gaagaattcc	ctgattggac	tactagatg	aagcaaaattg	ccaaattgtg	gagaaaagca	2220
agctcacaa	aaagagcacc	atatgtgcaa	aaagccagag	ataacagagc	tgctttacgc	2280
attaataaag	tacagatgtc	aaatgattcc	atgaaaaggc	agcaacagca	agatagcatt	2340
gatcccagct	ctcgtattga	ttcggagcct	tttaaagatc	ctttaaagca	aagagaatca	2400
gaacatgaac	aggaatggaa	atttagacag	caaatagcgtc	agaaaagtaa	gcagcaagct	2460
aaaattgaag	ccacacagaa	acttgaacag	gtgaaaaatg	agcagcagca	gcagcaacaa	2520
cagcaatttg	gttctcagca	tcttctgggtg	cagtctgggt	cagatacacc	aagtagtggg	2580
atacagagtc	ccttgacacc	tcagcctggc	aatggaaata	tgtctcctgc	acagtcattc	2640
cataaagaac	tgttttacaaa	acagccaccc	agtaccctta	cgtctacatc	ttcagatgat	2700
gtgtttgtaa	agccacaagc	tcacctcct	cctccagccc	catcccggt	ttccatccag	2760
gatagtcttt	ctcaggtcca	gacttctcag	ccaccctatg	cgcaagtgtt	ttcacctggg	2820
tcctctaatt	cagcaccacc	atctccaatg	gatccatag	caaaaattgtt	tggtaccctt	2880
cgaccacctc	ctgtgggcca	tagtttttcc	agaagaattt	ctgtgcacc	agtggaatac	2940
gtacacctt	tatcatcggt	atctaggccc	cttcaaatga	atgagacaac	agcaaatagg	3000
ccatcccctg	tcagagattt	atgttcttct	tcacagacaa	ataatgaccc	ctatgcaaaa	3060

cctccagaca	cacctaggcc	tgtgatgaca	gatcaatttc	ccaaatcctt	gggcctatcc	3120
cggtctcctg	tagtttcaga	acaaactgca	aaaggcccta	tagcagctgg	aaccagtgat	3180
cactttacta	aaccatctcc	tagggcagat	gtgtttcaaa	gacaaaggat	acctgactca	3240
tatgcacgac	ccttgttgac	acctgcacct	cttgatagtg	gtcctggacc	ttttaagact	3300
ccaatgcaac	ctcctccatc	ctctcaggat	ccttatggat	cagtgtcaca	ggcatcaagg	3360
cgattgtctg	ttgaccctta	tgaaaggcct	gctttgacac	caagacctat	agataatttt	3420
tctcataatc	agtcaaata	tccatatagt	cagcctcccc	ttaccccaca	tccagcagtg	3480
aatgaatctt	ttgcccaccc	ttcaagggtc	ttttcccagc	ctggaaccat	atcaaggcca	3540
acatctcagg	accatactc	ccaaccccca	ggaactccac	gacctgttgt	agattcttat	3600
tcccaatctt	caggaacagc	taggtccaat	acagaccctt	actctcaacc	tcctggaact	3660
ccccggccta	ctactgttga	cccatatagt	cagcagcccc	aaaccccaag	accatctaca	3720
caaactgact	tggtttgttac	acctgtaaca	aatcagaggc	attctgatcc	atatgctcat	3780
cctcctggaa	caccaagacc	tggaatttct	gtcccttact	ctcagccacc	agcaacacca	3840
aggccaagga	tttcagaggg	ttttactagg	tcctcaatga	caagaccagt	cctcatgcca	3900
aatcaggatc	ctttcctgca	agcagcaca	aaccgaggac	cagctttacc	tggcccgttg	3960
gtaaggccac	ctgatacatg	ttcccagaca	cctaggcccc	ctggacctgg	tctttcagac	4020
acatttagcc	gtgtttcccc	atctgctgcc	cgtgatccct	atgatcagtc	tccaatgact	4080
ccaagatctc	agtctgactc	ttttggaaca	agtcaaactg	cccatgatgt	tgctgatcag	4140
ccaaggcctg	gatcagaggg	gagcttctgt	gcatcttcaa	actctccaat	gcactcccaa	4200
ggccagcagt	tctctggtgt	ctcccaactt	cctggacctg	tgccaacttc	aggagtaact	4260
gatacacaga	atactgtaaa	tatggcccaa	gcagatacag	agaaattgag	acagcggcag	4320
aagttacgtg	aaatcattct	ccagcagcaa	cagcagaaga	agattgcagg	tcgacaggag	4380
aaggggtcac	aggactcacc	cgcagtgcct	catccagggc	ctcttcaaca	ctggcaacca	4440
gagaatgtta	accaggcttt	caccagacct	ccacctccct	atcctgggaa	cattaggtct	4500
cctgttgccc	ctccttttagg	acctagatat	gctgttttcc	caaaagatca	gcgtggacct	4560
tatcctcctg	atgttgctag	tatggggatg	agacctcatg	gatttagatt	tggatttcca	4620
ggaggtagtc	atggtaccat	gccgagtcaa	gagcgcttcc	ttgtgcctcc	tcagcaaata	4680
cagggatctg	gagtttctcc	acagctaaga	agatcagtat	ctgtagatat	gcctaggcct	4740
ttaaataact	cacaaatgaa	taatccagtt	ggacttcctc	agcatttttc	accacagagc	4800
ttgccagttc	agcagcacaa	catactgggc	caagcatata	ttgaactgag	acatagggct	4860
cctgacggaa	ggcaacggct	gcctttcagt	gctccacctg	gcagcgttgt	agaggcatct	4920
tctaacttga	gacatggaaa	cttcattccc	cggccagact	ttccgggccc	tagacacaca	4980
gaccccatgc	gacgacctcc	ccagggtcta	cctaatacagc	tacctgtgca	cccagatttg	5040
gaacaagtgc	caccatctca	acaagagcaa	ggtcattctg	tccattcatc	ttctatggtc	5100
atgaggactc	tgaacccatc	actagggtgg	gaatttttcag	aagctccttt	gtcaacatct	5160
gtaccgtctg	aaacaacgtc	tgataattta	cagataacca	cccagccttc	tgatggtcta	5220
gaggaaaaac	ttgattctga	tgacctttct	gtgaaggaa	tggatgttaa	agaccttgag	5280
ggggttgaag	tcaaaagactt	agatgatgaa	gatcttgaaa	acttaaattt	agatacagag	5340
gatggcaagg	tagttgaatt	ggatacttta	gataattttg	aaactaatga	tcccaacctg	5400
gatgacctct	taagggtcagg	agagtttgat	atcattgcat	atacagatcc	agaacttgac	5460
atgggagata	agaaaagcat	gtttaatgag	gaactagacc	ttccaattga	tgataagtta	5520
gataatcagt	gtgtatctgt	tgaacccaaa	aaaaaggaa	aagaaaacaa	aactctggtt	5580
ctctctgata	aacattccacc	acagaaaaaa	tccactgtta	ccaatgaggt	aaaaacggaa	5640
gtactgtctc	caaattctaa	ggtggaatcc	aaatgtgaaa	ctgaaaaaaa	tgatgagaat	5700
aaagataatg	ttgacactcc	ttgtcacacg	gcttctgctc	actcagacct	aaatgatgga	5760
gaaaagactt	ctttgcatcc	ttgtgatcca	gatctatttg	agaaaagaac	caatcgagaa	5820
actgctggcc	ccagtgcaca	tgtcattcag	gcattccactc	aactacctgc	tcaagatgta	5880
ataaactctt	gtggcataac	tggatcaact	ccagttctct	caagtttact	tgctaagtga	5940
aaatctgata	attcagacat	taggccatcg	gggtctccac	caccaccaac	tctgccggcc	6000
tccccatcca	atcatgtgtc	aagtttgcc	cctttcatag	caccgcctgg	ccgtgttttg	6060
gataatgcc	tgaattctaa	tgtgacagta	gtctctaggg	taaacctagt	tttttctcag	6120
ggtgtgcagg	taaacccagg	gctcattcca	ggtcaatcaa	cagttaacca	cagtctgggg	6180
acaggaaaac	ctgcaactca	aactgggect	caaacaagtc	agtctggtac	cagtagcatg	6240
tctggacccc	aacagctaat	gattcctcaa	acattagcac	agcagaatag	agagaggccc	6300

cttcttctag	aagaacagcc	tctacttcta	caggatcttt	tggatcaaga	aaggcaagaa	6360
cagcagcagc	aaagacagat	gcaagccatg	attcgtcagc	gatcagaacc	gttctttccct	6420
aatattgatt	ttgatgcaat	tacagatcct	ataatgaaag	ccaaaatggt	ggcccttaaa	6480
ggtataaata	aagtgatggc	acaaaacaat	ctgggcatgc	caccaatggt	gatgagcagg	6540
ttccctttta	tgggccaggt	ggtaactgga	acacagaaca	gtgaaggaca	gaaccttgga	6600
ccacaggcca	ttcctcagga	tggcagtata	acacatcaga	tttctaggcc	taatcctcca	6660
aattttggtc	caggctttgt	caatgattca	cagcgtaagc	agtatgaaga	gtggctccag	6720
gagacccaac	agctgcttca	aatgcagcag	aagtatcttg	aagaacaaat	tgggtgctcac	6780
agaaaaatcta	agaaggccct	ttcagctaaa	caacgtactg	ccaagaaaagc	tgggcgtgaa	6840
tttccagagg	aagatgcaga	acaactcaag	catgttactg	aacagcaaag	catggttcag	6900
aaacagctag	aacagattcg	taaacaacag	aaagaacatg	ctgaattgat	tgaagattat	6960
cggatcaaac	agcagcagca	atgtgcaatg	gccccaccta	ccatgatgcc	cagtgtccag	7020
ccccagccac	ccctaattcc	aggtgccact	ccaccacca	tgagccaacc	cacctttccc	7080
atggtgccac	agcagcttca	gcaccagcag	cacacaacag	ttattttctgg	ccatactagc	7140
cctgttagaa	tggccagttt	acctggatgg	caacccaaca	gtgctcctgc	ccacctgccc	7200
ctcaatcctc	ctagaattca	gcccccaatt	gccagttac	caataaaaaac	ttgtacacca	7260
gccccagggg	cagtctcaaa	tgcaaatcca	cagagtggac	caccacctcg	ggtagaattt	7320
gatgacaaca	atcccttttag	tgaaagtttt	caagaacggg	aacgtaagga	acgtttacga	7380
gaacagcaag	agagacaacg	gatccaactc	atgcaggagg	tagatagaca	aagagctttg	7440
cagcagagga	tggaaatgga	gcagcatggt	atggtgggct	ctgagataag	tagtagtagg	7500
acatctgtgt	cccagattcc	cttctacagt	tccgacttac	cttgtgattt	tatgcaacct	7560
ctaggacccc	ttcagcagtc	tccacaacac	caacagcaaa	tggggcaggt	tttacagcag	7620
cagaatatac	aacaaggatc	aattaattca	ccctccaccc	aaactttcat	gcagactaat	7680
gagcgaaggc	aggtagggcc	tccttcattt	gttcttgatt	caccatcaat	ccctgttgga	7740
agcccaaatt	tttcttctgt	gaagcagggg	catggaaatc	tttctgggac	cagcttccag	7800
cagtccccag	tgaggccctc	ttttacacct	gctttaccag	cagcacctcc	agtagctaat	7860
agcagtctcc	catgtggcca	agatttctact	ataacccatg	gacacagtta	tccgggatca	7920
acccaatcgc	tcattcagtt	gtatttctgat	ataatcccag	aggaaaaagg	gaaaaagaaa	7980
agaacaagaa	agaagaaaag	agatgatgat	gcagaatcca	ccaaggctcc	atcaactccc	8040
cattcagata	taactgcccc	accgactcca	ggcatctcag	aaactacctc	tactcctgca	8100
gtgagcacac	ccagtgaagt	tcctcaacaa	gccgaccaag	agtcggtgga	accagtcggc	8160
ccatccactc	ccaatatggc	agcaggccag	ctatgtacag	aattagagaa	caaactgccc	8220
aatagtgatt	tctcacaagc	aactccaaat	caacagacgt	atgcaaattc	agaagtagac	8280
aagctctcca	tggaaacccc	tgccaaaaca	gaagagataa	aactggaaaa	ggctgagaca	8340
gagtcttgcc	caggccaaga	ggagccataa	ttggagggaac	agaatggtag	taaggtagaa	8400
ggaaacgctg	tagcctgtcc	tgtctectca	gcacagagtc	ctccccattc	tgctggggcc	8460
cctgctgcca	aaggagactc	agggaatgaa	cttctgaaac	acttgttgaa	aaataaaaag	8520
tcactctctc	ttttgaatca	aaaacctgag	ggcagtattt	gttcagaaga	tgactgtaca	8580
aaggataata	aactagtttg	gaagcagaac	ccagctgaag	gactgcaaac	tttgggggct	8640
caaatgcaag	gtggttttgg	atgtggcaac	cagttgccaa	aaacagatgg	aggaagtga	8700
accaagaaac	agcgaagcaa	acggactcag	aggacgggtg	agaaagcagc	acctcgctca	8760
aagaaaaagga	aaaaggacga	agaggagaaa	caagctatgt	actctagcac	tgacacgttt	8820
acccacttga	aacaggtgag	gcagctctct	ctgctccctc	taatggaacc	aatcattgga	8880
gtgaactttg	cgcactttct	tccttatggc	agtggccaat	ttaatagtgg	gaatcgactt	8940
ctaggaactt	ttggcagtg	taccctggaa	ggggtttcgg	actactattc	tcagttgatc	9000
tacaagcaga	ataatttaag	taatcctcca	acaccccctg	cctctcttcc	tcctacacca	9060
cctcctatgg	cttgtcagaa	gatggccaat	ggttttgcaa	caactgaaga	acttgctgga	9120
aaagccggag	tgttagtgag	ccatgaagtt	acaaaaactc	taggacctaa	accatttcag	9180
ctgcccttca	gaccccagga	cgacttggtg	gcccagagctc	ttgctcaggg	ccccagaca	9240
gttgatgtgc	cagcctccct	cccaacacca	cctcataaca	atcaggaaga	attaaggata	9300
caggatcact	gtggtgatcg	agataactcct	gacagttttg	ttccctcatc	ctctcctgag	9360
agtgtggttg	gggtagaagt	gagcaggtat	ccagatctgt	cattgggtcaa	ggaggagcct	9420
ccagaaccgg	tgcctgcccc	catcattcca	attcttcccta	gcactgctgg	gaaaagtcca	9480
gaatcaagaa	ggaatgacat	caaaaactgag	ccaggcactt	tatatatttg	gtcacctttt	9540

<400> 423

```
<210> 424
<211> 1549
<212> DNA
<213> Homo sapiens
```

```
<210> 425
<211> 4019
<212> PRT
<213> Homo sapiens
```

Val Cys Gly Ser Phe Gly Gln Gly Ala Glu Gly Arg Leu Leu Ala Cys

Ser	Gln	Cys	Gly	Gln	Cys	Tyr	His	Pro	Tyr	Cys	Val	Ser	Ile	Lys	Ile
		35					40					45			
Thr	Lys	Val	Val	Leu	Ser	Lys	Gly	Trp	Arg	Cys	Leu	Glu	Cys	Thr	Val
	50					55					60				
Cys	Glu	Ala	Cys	Gly	Lys	Ala	Thr	Asp	Pro	Gly	Arg	Leu	Leu	Leu	Cys
65					70					75					80
Asp	Asp	Cys	Asp	Ile	Ser	Tyr	His	Thr	Tyr	Cys	Leu	Asp	Pro	Pro	Leu
				85					90					95	
Gln	Thr	Val	Pro	Lys	Gly	Gly	Trp	Lys	Cys	Lys	Trp	Cys	Val	Trp	Cys
			100					105					110		
Arg	His	Cys	Gly	Ala	Thr	Ser	Ala	Gly	Leu	Arg	Cys	Glu	Trp	Gln	Asn
		115					120					125			
Asn	Tyr	Thr	Gln	Cys	Ala	Pro	Cys	Ala	Ser	Leu	Ser	Ser	Cys	Pro	Val
	130					135					140				
Cys	Tyr	Arg	Asn	Tyr	Arg	Glu	Glu	Asp	Leu	Ile	Leu	Gln	Cys	Arg	Gln
145					150					155					160
Cys	Asp	Arg	Trp	Met	His	Ala	Val	Cys	Gln	Asn	Leu	Asn	Thr	Glu	Glu
				165					170					175	
Glu	Val	Glu	Asn	Val	Ala	Asp	Ile	Gly	Phe	Asp	Cys	Ser	Met	Cys	Arg
			180					185					190		
Pro	Tyr	Met	Pro	Ala	Ser	Asn	Val	Pro	Ser	Ser	Asp	Cys	Cys	Glu	Ser
		195					200					205			
Ser	Leu	Val	Ala	Gln	Ile	Val	Thr	Lys	Val	Lys	Glu	Leu	Asp	Pro	Pro
	210					215					220				
Lys	Thr	Tyr	Thr	Gln	Asp	Gly	Val	Cys	Leu	Thr	Glu	Ser	Gly	Met	Thr
225					230					235					240
Gln	Leu	Gln	Ser	Leu	Thr	Val	Thr	Val	Pro	Arg	Arg	Lys	Arg	Ser	Lys
				245					250					255	
Pro	Lys	Leu	Lys	Leu	Lys	Ile	Ile	Asn	Gln	Asn	Ser	Val	Ala	Val	Leu
			260					265					270		
Gln	Thr	Pro	Pro	Asp	Ile	Gln	Ser	Glu	His	Ser	Arg	Asp	Gly	Glu	Met
		275					280					285			
Asp	Asp	Ser	Arg	Glu	Gly	Glu	Leu	Met	Asp	Cys	Asp	Gly	Lys	Ser	Glu
	290					295					300				
Ser	Ser	Pro	Glu	Arg	Glu	Ala	Val	Asp	Asp	Glu	Thr	Lys	Gly	Val	Glu

305				310				315				320			
Gly	Thr	Asp	Gly	Val	Lys	Lys	Arg	Lys	Arg	Lys	Pro	Tyr	Arg	Pro	Gly
325				330				335							
Ile	Gly	Gly	Phe	Met	Val	Arg	Gln	Arg	Ser	Arg	Thr	Gly	Gln	Gly	Lys
340				345				350							
Thr	Lys	Arg	Ser	Val	Ile	Arg	Lys	Asp	Ser	Ser	Gly	Ser	Ile	Ser	Glu
355				360				365							
Gln	Leu	Pro	Cys	Arg	Asp	Asp	Gly	Trp	Ser	Glu	Gln	Leu	Pro	Asp	Thr
370				375				380							
Leu	Val	Asp	Glu	Ser	Val	Ser	Val	Thr	Glu	Ser	Thr	Glu	Lys	Ile	Lys
385				390				395				400			
Lys	Arg	Tyr	Arg	Lys	Arg	Lys	Asn	Lys	Leu	Glu	Glu	Thr	Phe	Pro	Ala
405				410				415							
Tyr	Leu	Gln	Glu	Ala	Phe	Phe	Gly	Lys	Asp	Leu	Leu	Asp	Thr	Ser	Arg
420				425				430							
Gln	Ser	Lys	Ile	Ser	Leu	Asp	Asn	Leu	Ser	Glu	Asp	Gly	Ala	Gln	Leu
435				440				445							
Leu	Tyr	Lys	Thr	Asn	Met	Asn	Thr	Gly	Phe	Leu	Asp	Pro	Ser	Leu	Asp
450				455				460							
Pro	Leu	Leu	Ser	Ser	Ser	Ser	Ala	Pro	Thr	Lys	Ser	Gly	Thr	His	Gly
465				470				475				480			
Pro	Ala	Asp	Asp	Pro	Leu	Ala	Asp	Ile	Ser	Glu	Val	Leu	Asn	Thr	Asp
485				490				495							
Asp	Asp	Ile	Leu	Gly	Ile	Ile	Ser	Asp	Asp	Leu	Ala	Lys	Ser	Val	Asp
500				505				510							
His	Ser	Asp	Ile	Gly	Pro	Val	Thr	Asp	Asp	Pro	Ser	Ser	Leu	Pro	Gln
515				520				525							
Pro	Asn	Val	Asn	Gln	Ser	Ser	Arg	Pro	Leu	Ser	Glu	Glu	Gln	Leu	Asp
530				535				540							
Gly	Ile	Leu	Ser	Pro	Glu	Leu	Asp	Lys	Met	Val	Thr	Asp	Gly	Ala	Ile
545				550				555				560			
Leu	Gly	Lys	Leu	Tyr	Lys	Ile	Pro	Glu	Leu	Gly	Gly	Lys	Asp	Val	Glu
565				570				575							
Asp	Leu	Phe	Thr	Ala	Val	Leu	Ser	Pro	Ala	Asn	Thr	Gln	Pro	Thr	Pro
580				585				590							
Leu	Pro	Gln	Pro	Pro	Pro	Pro	Thr	Gln	Leu	Leu	Pro	Ile	His	Asn	Gln

595					600					605					
Asp	Ala	Phe	Ser	Arg	Met	Pro	Leu	Met	Asn	Gly	Leu	Ile	Gly	Ser	Ser
610						615					620				
Pro	His	Leu	Pro	His	Asn	Ser	Leu	Pro	Pro	Gly	Ser	Gly	Leu	Gly	Thr
625					630					635					640
Phe	Ser	Ala	Ile	Ala	Gln	Ser	Ser	Tyr	Pro	Asp	Ala	Arg	Asp	Lys	Asn
				645					650					655	
Ser	Ala	Phe	Asn	Pro	Met	Ala	Ser	Asp	Pro	Asn	Asn	Ser	Trp	Thr	Ser
			660					665					670		
Ser	Ala	Pro	Thr	Val	Glu	Gly	Glu	Asn	Asp	Thr	Met	Ser	Asn	Ala	Gln
		675					680					685			
Arg	Ser	Thr	Leu	Lys	Trp	Glu	Lys	Glu	Glu	Ala	Leu	Gly	Glu	Met	Ala
690						695					700				
Thr	Val	Ala	Pro	Val	Leu	Tyr	Thr	Asn	Ile	Asn	Phe	Pro	Asn	Leu	Lys
705					710					715					720
Glu	Glu	Phe	Pro	Asp	Trp	Thr	Thr	Arg	Val	Lys	Gln	Ile	Ala	Lys	Leu
			725						730					735	
Trp	Arg	Lys	Ala	Ser	Ser	Gln	Glu	Arg	Ala	Pro	Tyr	Val	Gln	Lys	Ala
			740					745					750		
Arg	Asp	Asn	Arg	Ala	Ala	Leu	Arg	Ile	Asn	Lys	Val	Gln	Met	Ser	Asn
		755					760						765		
Asp	Ser	Met	Lys	Arg	Gln	Gln	Gln	Gln	Asp	Ser	Ile	Asp	Pro	Ser	Ser
		770				775					780				
Arg	Ile	Asp	Ser	Glu	Leu	Phe	Lys	Asp	Pro	Leu	Lys	Gln	Arg	Glu	Ser
785				790					795					800	
Glu	His	Glu	Gln	Glu	Trp	Lys	Phe	Arg	Gln	Gln	Met	Arg	Gln	Lys	Ser
			805						810				815		
Lys	Gln	Gln	Ala	Lys	Ile	Glu	Ala	Thr	Gln	Lys	Leu	Glu	Gln	Val	Lys
			820					825					830		
Asn	Glu	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Phe	Gly	Ser	Gln	His	Leu
		835					840					845			
Leu	Val	Gln	Ser	Gly	Ser	Asp	Thr	Pro	Ser	Ser	Gly	Ile	Gln	Ser	Pro
	850					855					860				
Leu	Thr	Pro	Gln	Pro	Gly	Asn	Gly	Asn	Met	Ser	Pro	Ala	Gln	Ser	Phe
865				870					875					880	
His	Lys	Glu	Leu	Phe	Thr	Lys	Gln	Pro	Pro	Ser	Thr	Pro	Thr	Ser	Thr

0966770-092000

885																890						895					
Ser	Ser	Asp	Asp	Val	Phe	Val	Lys	Pro	Gln	Ala	Pro	Pro	Pro	Pro	Pro												
			900					905							910												
Ala	Pro	Ser	Arg	Ile	Pro	Ile	Gln	Asp	Ser	Leu	Ser	Gln	Ala	Gln	Thr												
		915					920							925													
Ser	Gln	Pro	Pro	Ser	Pro	Gln	Val	Phe	Ser	Pro	Gly	Ser	Ser	Asn	Ser												
		930					935							940													
Arg	Pro	Pro	Ser	Pro	Met	Asp	Pro	Tyr	Ala	Lys	Met	Val	Gly	Thr	Pro												
		945				950							955			960											
Arg	Pro	Pro	Pro	Val	Gly	His	Ser	Phe	Ser	Arg	Arg	Asn	Ser	Ala	Ala												
				965					970							975											
Pro	Val	Glu	Asn	Cys	Thr	Pro	Leu	Ser	Ser	Val	Ser	Arg	Pro	Leu	Gln												
			980					985							990												
Met	Asn	Glu	Thr	Thr	Ala	Asn	Arg	Pro	Ser	Pro	Val	Arg	Asp	Leu	Cys												
		995					1000							1005													
Ser	Ser	Ser	Thr	Thr	Asn	Asn	Asp	Pro	Tyr	Ala	Lys	Pro	Pro	Asp	Thr												
		1010					1015							1020													
Pro	Arg	Pro	Val	Met	Thr	Asp	Gln	Phe	Pro	Lys	Ser	Leu	Gly	Leu	Ser												
		1025				1030							1035			1040											
Arg	Ser	Pro	Val	Val	Ser	Glu	Gln	Thr	Ala	Lys	Gly	Pro	Ile	Ala	Ala												
				1045					1050							1055											
Gly	Thr	Ser	Asp	His	Phe	Thr	Lys	Pro	Ser	Pro	Arg	Ala	Asp	Val	Phe												
			1060					1065							1070												
Gln	Arg	Gln	Arg	Ile	Pro	Asp	Ser	Tyr	Ala	Arg	Pro	Leu	Leu	Thr	Pro												
		1075					1080							1085													
Ala	Pro	Leu	Asp	Ser	Gly	Pro	Gly	Pro	Phe	Lys	Thr	Pro	Met	Gln	Pro												
		1090					1095							1100													
Pro	Pro	Ser	Ser	Gln	Asp	Pro	Tyr	Gly	Ser	Val	Ser	Gln	Ala	Ser	Arg												
		1105				1110							1115			1120											
Arg	Leu	Ser	Val	Asp	Pro	Tyr	Glu	Arg	Pro	Ala	Leu	Thr	Pro	Arg	Pro												
				1125					1130							1135											
Ile	Asp	Asn	Phe	Ser	His	Asn	Gln	Ser	Asn	Asp	Pro	Tyr	Ser	Gln	Pro												
			1140					1145							1150												
Pro	Leu	Thr	Pro	His	Pro	Ala	Val	Asn	Glu	Ser	Phe	Ala	His	Pro	Ser												
		1155					1160							1165													
Arg	Ala	Phe	Ser	Gln	Pro	Gly	Thr	Ile	Ser	Arg	Pro	Thr	Ser	Gln	Asp												

1170 1175 1180
 Pro Tyr Ser Gln Pro Pro Gly Thr Pro Arg Pro Val Val Asp Ser Tyr
 1185 1190 1195 1200
 Ser Gln Ser Ser Gly Thr Ala Arg Ser Asn Thr Asp Pro Tyr Ser Gln
 1205 1210 1215
 Pro Pro Gly Thr Pro Arg Pro Thr Thr Val Asp Pro Tyr Ser Gln Gln
 1220 1225 1230
 Pro Gln Thr Pro Arg Pro Ser Thr Gln Thr Asp Leu Phe Val Thr Pro
 1235 1240 1245
 Val Thr Asn Gln Arg His Ser Asp Pro Tyr Ala His Pro Pro Gly Thr
 1250 1255 1260
 Pro Arg Pro Gly Ile Ser Val Pro Tyr Ser Gln Pro Pro Ala Thr Pro
 1265 1270 1275 1280
 Arg Pro Arg Ile Ser Glu Gly Phe Thr Arg Ser Ser Met Thr Arg Pro
 1285 1290 1295
 Val Leu Met Pro Asn Gln Asp Pro Phe Leu Gln Ala Ala Gln Asn Arg
 1300 1305 1310
 Gly Pro Ala Leu Pro Gly Pro Leu Val Arg Pro Pro Asp Thr Cys Ser
 1315 1320 1325
 Gln Thr Pro Arg Pro Pro Gly Pro Gly Leu Ser Asp Thr Phe Ser Arg
 1330 1335 1340
 Val Ser Pro Ser Ala Ala Arg Asp Pro Tyr Asp Gln Ser Pro Met Thr
 1345 1350 1355 1360
 Pro Arg Ser Gln Ser Asp Ser Phe Gly Thr Ser Gln Thr Ala His Asp
 1365 1370 1375
 Val Ala Asp Gln Pro Arg Pro Gly Ser Glu Gly Ser Phe Cys Ala Ser
 1380 1385 1390
 Ser Asn Ser Pro Met His Ser Gln Gly Gln Gln Phe Ser Gly Val Ser
 1395 1400 1405
 Gln Leu Pro Gly Pro Val Pro Thr Ser Gly Val Thr Asp Thr Gln Asn
 1410 1415 1420
 Thr Val Asn Met Ala Gln Ala Asp Thr Glu Lys Leu Arg Gln Arg Gln
 1425 1430 1435 1440
 Lys Leu Arg Glu Ile Ile Leu Gln Gln Gln Gln Lys Lys Ile Ala
 1445 1450 1455
 Gly Arg Gln Glu Lys Gly Ser Gln Asp Ser Pro Ala Val Pro His Pro

00667170-092000

1460					1465					1470						
Gly	Pro	Leu	Gln	His	Trp	Gln	Pro	Glu	Asn	Val	Asn	Gln	Ala	Phe	Thr	
1475					1480					1485						
Arg	Pro	Pro	Pro	Pro	Tyr	Pro	Gly	Asn	Ile	Arg	Ser	Pro	Val	Ala	Pro	
1490					1495					1500						
Pro	Leu	Gly	Pro	Arg	Tyr	Ala	Val	Phe	Pro	Lys	Asp	Gln	Arg	Gly	Pro	
1505					1510					1515					1520	
Tyr	Pro	Pro	Asp	Val	Ala	Ser	Met	Gly	Met	Arg	Pro	His	Gly	Phe	Arg	
1525					1530					1535						
Phe	Gly	Phe	Pro	Gly	Gly	Ser	His	Gly	Thr	Met	Pro	Ser	Gln	Glu	Arg	
1540					1545					1550						
Phe	Leu	Val	Pro	Pro	Gln	Gln	Ile	Gln	Gly	Ser	Gly	Val	Ser	Pro	Gln	
1555					1560					1565						
Leu	Arg	Arg	Ser	Val	Ser	Val	Asp	Met	Pro	Arg	Pro	Leu	Asn	Asn	Ser	
1570					1575					1580						
Gln	Met	Asn	Asn	Pro	Val	Gly	Leu	Pro	Gln	His	Phe	Ser	Pro	Gln	Ser	
1585					1590					1595					1600	
Leu	Pro	Val	Gln	Gln	His	Asn	Ile	Leu	Gly	Gln	Ala	Tyr	Ile	Glu	Leu	
1605					1610					1615						
Arg	His	Arg	Ala	Pro	Asp	Gly	Arg	Gln	Arg	Leu	Pro	Phe	Ser	Ala	Pro	
1620					1625					1630						
Pro	Gly	Ser	Val	Val	Glu	Ala	Ser	Ser	Asn	Leu	Arg	His	Gly	Asn	Phe	
1635					1640					1645						
Ile	Pro	Arg	Pro	Asp	Phe	Pro	Gly	Pro	Arg	His	Thr	Asp	Pro	Met	Arg	
1650					1655					1660						
Arg	Pro	Pro	Gln	Gly	Leu	Pro	Asn	Gln	Leu	Pro	Val	His	Pro	Asp	Leu	
1665					1670					1675					1680	
Glu	Gln	Val	Pro	Pro	Ser	Gln	Gln	Glu	Gln	Gly	His	Ser	Val	His	Ser	
1685					1690					1695						
Ser	Ser	Met	Val	Met	Arg	Thr	Leu	Asn	His	Pro	Leu	Gly	Gly	Glu	Phe	
1700					1705					1710						
Ser	Glu	Ala	Pro	Leu	Ser	Thr	Ser	Val	Pro	Ser	Glu	Thr	Thr	Ser	Asp	
1715					1720					1725						
Asn	Leu	Gln	Ile	Thr	Thr	Gln	Pro	Ser	Asp	Gly	Leu	Glu	Glu	Lys	Leu	
1730					1735					1740						
Asp	Ser	Asp	Asp	Pro	Ser	Val	Lys	Glu	Leu	Asp	Val	Lys	Asp	Leu	Glu	

1745											1750											1755											1760
Gly	Val	Glu	Val	Lys	Asp	Leu	Asp	Asp	Glu	Asp	Leu	Glu	Asn	Leu	Asn																		
				1765								1770								1775													
Leu	Asp	Thr	Glu	Asp	Gly	Lys	Val	Val	Glu	Leu	Asp	Thr	Leu	Asp	Asn																		
				1780								1785								1790													
Leu	Glu	Thr	Asn	Asp	Pro	Asn	Leu	Asp	Asp	Leu	Leu	Arg	Ser	Gly	Glu																		
				1795								1800								1805													
Phe	Asp	Ile	Ile	Ala	Tyr	Thr	Asp	Pro	Glu	Leu	Asp	Met	Gly	Asp	Lys																		
				1810								1815								1820													
Lys	Ser	Met	Phe	Asn	Glu	Glu	Leu	Asp	Leu	Pro	Ile	Asp	Asp	Lys	Leu																		
1825								1830								1835								1840									
Asp	Asn	Gln	Cys	Val	Ser	Val	Glu	Pro	Lys	Lys	Lys	Glu	Gln	Glu	Asn																		
				1845								1850								1855													
Lys	Thr	Leu	Val	Leu	Ser	Asp	Lys	His	Ser	Pro	Gln	Lys	Lys	Ser	Thr																		
				1860								1865								1870													
Val	Thr	Asn	Glu	Val	Lys	Thr	Glu	Val	Leu	Ser	Pro	Asn	Ser	Lys	Val																		
				1875								1880								1885													
Glu	Ser	Lys	Cys	Glu	Thr	Glu	Lys	Asn	Asp	Glu	Asn	Lys	Asp	Asn	Val																		
				1890								1895								1900													
Asp	Thr	Pro	Cys	Ser	Gln	Ala	Ser	Ala	His	Ser	Asp	Leu	Asn	Asp	Gly																		
1905								1910								1915								1920									
Glu	Lys	Thr	Ser	Leu	His	Pro	Cys	Asp	Pro	Asp	Leu	Phe	Glu	Lys	Arg																		
				1925								1930								1935													
Thr	Asn	Arg	Glu	Thr	Ala	Gly	Pro	Ser	Ala	Asn	Val	Ile	Gln	Ala	Ser																		
				1940								1945								1950													
Thr	Gln	Leu	Pro	Ala	Gln	Asp	Val	Ile	Asn	Ser	Cys	Gly	Ile	Thr	Gly																		
				1955								1960								1965													
Ser	Thr	Pro	Val	Leu	Ser	Ser	Leu	Leu	Ala	Asn	Glu	Lys	Ser	Asp	Asn																		
				1970								1975								1980													
Ser	Asp	Ile	Arg	Pro	Ser	Gly	Ser	Pro	Pro	Pro	Pro	Thr	Leu	Pro	Ala																		
1985								1990								1995								2000									
Ser	Pro	Ser	Asn	His	Val	Ser	Ser	Leu	Pro	Pro	Phe	Ile	Ala	Pro	Pro																		
				2005								2010								2015													
Gly	Arg	Val	Leu	Asp	Asn	Ala	Met	Asn	Ser	Asn	Val	Thr	Val	Val	Ser																		
				2020								2025								2030													
Arg	Val	Asn	His	Val	Phe	Ser	Gln	Gly	Val	Gln	Val	Asn	Pro	Gly	Leu																		

2035	2040	2045
Ile Pro Gly Gln Ser Thr Val Asn His Ser Leu Gly Thr Gly Lys Pro		
2050	2055	2060
Ala Thr Gln Thr Gly Pro Gln Thr Ser Gln Ser Gly Thr Ser Ser Met		
2065	2070	2075 2080
Ser Gly Pro Gln Gln Leu Met Ile Pro Gln Thr Leu Ala Gln Gln Asn		
2085	2090	2095
Arg Glu Arg Pro Leu Leu Leu Glu Glu Gln Pro Leu Leu Leu Gln Asp		
2100	2105	2110
Leu Leu Asp Gln Glu Arg Gln Glu Gln Gln Gln Arg Gln Met Gln		
2115	2120	2125
Ala Met Ile Arg Gln Arg Ser Glu Pro Phe Phe Pro Asn Ile Asp Phe		
2130	2135	2140
Asp Ala Ile Thr Asp Pro Ile Met Lys Ala Lys Met Val Ala Leu Lys		
2145	2150	2155 2160
Gly Ile Asn Lys Val Met Ala Gln Asn Asn Leu Gly Met Pro Pro Met		
2165	2170	2175
Val Met Ser Arg Phe Pro Phe Met Gly Gln Val Val Thr Gly Thr Gln		
2180	2185	2190
Asn Ser Glu Gly Gln Asn Leu Gly Pro Gln Ala Ile Pro Gln Asp Gly		
2195	2200	2205
Ser Ile Thr His Gln Ile Ser Arg Pro Asn Pro Pro Asn Phe Gly Pro		
2210	2215	2220
Gly Phe Val Asn Asp Ser Gln Arg Lys Gln Tyr Glu Glu Trp Leu Gln		
2225	2230	2235 2240
Glu Thr Gln Gln Leu Leu Gln Met Gln Gln Lys Tyr Leu Glu Glu Gln		
2245	2250	2255
Ile Gly Ala His Arg Lys Ser Lys Lys Ala Leu Ser Ala Lys Gln Arg		
2260	2265	2270
Thr Ala Lys Lys Ala Gly Arg Glu Phe Pro Glu Glu Asp Ala Glu Gln		
2275	2280	2285
Leu Lys His Val Thr Glu Gln Gln Ser Met Val Gln Lys Gln Leu Glu		
2290	2295	2300
Gln Ile Arg Lys Gln Gln Lys Glu His Ala Glu Leu Ile Glu Asp Tyr		
2305	2310	2315 2320
Arg Ile Lys Gln Gln Gln Gln Cys Ala Met Ala Pro Pro Thr Met Met		

000260"02'2955

2325	2330	2335
Pro Ser Val Gln Pro Gln Pro Pro Leu Ile Pro Gly Ala Thr Pro Pro 2340 2345 2350		
Thr Met Ser Gln Pro Thr Phe Pro Met Val Pro Gln Gln Leu Gln His 2355 2360 2365		
Gln Gln His Thr Thr Val Ile Ser Gly His Thr Ser Pro Val Arg Met 2370 2375 2380		
Pro Ser Leu Pro Gly Trp Gln Pro Asn Ser Ala Pro Ala His Leu Pro 2385 2390 2395 2400		
Leu Asn Pro Pro Arg Ile Gln Pro Pro Ile Ala Gln Leu Pro Ile Lys 2405 2410 2415		
Thr Cys Thr Pro Ala Pro Gly Thr Val Ser Asn Ala Asn Pro Gln Ser 2420 2425 2430		
Gly Pro Pro Pro Arg Val Glu Phe Asp Asp Asn Asn Pro Phe Ser Glu 2435 2440 2445		
Ser Phe Gln Glu Arg Glu Arg Lys Glu Arg Leu Arg Glu Gln Gln Glu 2450 2455 2460		
Arg Gln Arg Ile Gln Leu Met Gln Glu Val Asp Arg Gln Arg Ala Leu 2465 2470 2475 2480		
Gln Gln Arg Met Glu Met Glu Gln His Gly Met Val Gly Ser Glu Ile 2485 2490 2495		
Ser Ser Ser Arg Thr Ser Val Ser Gln Ile Pro Phe Tyr Ser Ser Asp 2500 2505 2510		
Leu Pro Cys Asp Phe Met Gln Pro Leu Gly Pro Leu Gln Gln Ser Pro 2515 2520 2525		
Gln His Gln Gln Gln Met Gly Gln Val Leu Gln Gln Gln Asn Ile Gln 2530 2535 2540		
Gln Gly Ser Ile Asn Ser Pro Ser Thr Gln Thr Phe Met Gln Thr Asn 2545 2550 2555 2560		
Glu Arg Arg Gln Val Gly Pro Pro Ser Phe Val Pro Asp Ser Pro Ser 2565 2570 2575		
Ile Pro Val Gly Ser Pro Asn Phe Ser Ser Val Lys Gln Gly His Gly 2580 2585 2590		
Asn Leu Ser Gly Thr Ser Phe Gln Gln Ser Pro Val Arg Pro Ser Phe 2595 2600 2605		
Thr Pro Ala Leu Pro Ala Ala Pro Pro Val Ala Asn Ser Ser Leu Pro		

2610						2615						2620					
Cys	Gly	Gln	Asp	Ser	Thr	Ile	Thr	His	Gly	His	Ser	Tyr	Pro	Gly	Ser		
2625							2630				2635				2640		
Thr	Gln	Ser	Leu	Ile	Gln	Leu	Tyr	Ser	Asp	Ile	Ile	Pro	Glu	Glu	Lys		
				2645					2650					2655			
Gly	Lys	Lys	Lys	Arg	Thr	Arg	Lys	Lys	Lys	Arg	Asp	Asp	Asp	Ala	Glu		
			2660					2665					2670				
Ser	Thr	Lys	Ala	Pro	Ser	Thr	Pro	His	Ser	Asp	Ile	Thr	Ala	Pro	Pro		
		2675					2680					2685					
Thr	Pro	Gly	Ile	Ser	Glu	Thr	Thr	Ser	Thr	Pro	Ala	Val	Ser	Thr	Pro		
	2690					2695					2700						
Ser	Glu	Leu	Pro	Gln	Gln	Ala	Asp	Gln	Glu	Ser	Val	Glu	Pro	Val	Gly		
2705					2710					2715					2720		
Pro	Ser	Thr	Pro	Asn	Met	Ala	Ala	Gly	Gln	Leu	Cys	Thr	Glu	Leu	Glu		
				2725					2730					2735			
Asn	Lys	Leu	Pro	Asn	Ser	Asp	Phe	Ser	Gln	Ala	Thr	Pro	Asn	Gln	Gln		
			2740					2745					2750				
Thr	Tyr	Ala	Asn	Ser	Glu	Val	Asp	Lys	Leu	Ser	Met	Glu	Thr	Pro	Ala		
	2755						2760					2765					
Lys	Thr	Glu	Glu	Ile	Lys	Leu	Glu	Lys	Ala	Glu	Thr	Glu	Ser	Cys	Pro		
	2770					2775					2780						
Gly	Gln	Glu	Glu	Pro	Lys	Leu	Glu	Glu	Gln	Asn	Gly	Ser	Lys	Val	Glu		
2785					2790					2795					2800		
Gly	Asn	Ala	Val	Ala	Cys	Pro	Val	Ser	Ser	Ala	Gln	Ser	Pro	Pro	His		
				2805					2810					2815			
Ser	Ala	Gly	Ala	Pro	Ala	Ala	Lys	Gly	Asp	Ser	Gly	Asn	Glu	Leu	Leu		
		2820						2825					2830				
Lys	His	Leu	Leu	Lys	Asn	Lys	Lys	Ser	Ser	Ser	Leu	Leu	Asn	Gln	Lys		
	2835					2840						2845					
Pro	Glu	Gly	Ser	Ile	Cys	Ser	Glu	Asp	Asp	Cys	Thr	Lys	Asp	Asn	Lys		
	2850					2855					2860						
Leu	Val	Glu	Lys	Gln	Asn	Pro	Ala	Glu	Gly	Leu	Gln	Thr	Leu	Gly	Ala		
2865					2870					2875					2880		
Gln	Met	Gln	Gly	Gly	Phe	Gly	Cys	Gly	Asn	Gln	Leu	Pro	Lys	Thr	Asp		
			2885						2890					2895			
Gly	Gly	Ser	Glu	Thr	Lys	Lys	Gln	Arg	Ser	Lys	Arg	Thr	Gln	Arg	Thr		

2900					2905					2910						
Gly	Glu	Lys	Ala	Ala	Pro	Arg	Ser	Lys	Lys	Arg	Lys	Lys	Asp	Glu	Glu	
		2915					2920					2925				
Glu	Lys	Gln	Ala	Met	Tyr	Ser	Ser	Thr	Asp	Thr	Phe	Thr	His	Leu	Lys	
		2930				2935					2940					
Gln	Val	Arg	Gln	Leu	Ser	Leu	Leu	Pro	Leu	Met	Glu	Pro	Ile	Ile	Gly	
		2945			2950					2955						2960
Val	Asn	Phe	Ala	His	Phe	Leu	Pro	Tyr	Gly	Ser	Gly	Gln	Phe	Asn	Ser	
				2965					2970					2975		
Gly	Asn	Arg	Leu	Leu	Gly	Thr	Phe	Gly	Ser	Ala	Thr	Leu	Glu	Gly	Val	
			2980					2985					2990			
Ser	Asp	Tyr	Tyr	Ser	Gln	Leu	Ile	Tyr	Lys	Gln	Asn	Asn	Leu	Ser	Asn	
		2995					3000					3005				
Pro	Pro	Thr	Pro	Pro	Ala	Ser	Leu	Pro	Pro	Thr	Pro	Pro	Pro	Met	Ala	
		3010				3015					3020					
Cys	Gln	Lys	Met	Ala	Asn	Gly	Phe	Ala	Thr	Thr	Glu	Glu	Leu	Ala	Gly	
		3025			3030					3035						3040
Lys	Ala	Gly	Val	Leu	Val	Ser	His	Glu	Val	Thr	Lys	Thr	Leu	Gly	Pro	
				3045					3050					3055		
Lys	Pro	Phe	Gln	Leu	Pro	Phe	Arg	Pro	Gln	Asp	Asp	Leu	Leu	Ala	Arg	
			3060					3065					3070			
Ala	Leu	Ala	Gln	Gly	Pro	Lys	Thr	Val	Asp	Val	Pro	Ala	Ser	Leu	Pro	
		3075					3080					3085				
Thr	Pro	Pro	His	Asn	Asn	Gln	Glu	Glu	Leu	Arg	Ile	Gln	Asp	His	Cys	
		3090				3095					3100					
Gly	Asp	Arg	Asp	Thr	Pro	Asp	Ser	Phe	Val	Pro	Ser	Ser	Ser	Pro	Glu	
		3105			3110					3115						3120
Ser	Val	Val	Gly	Val	Glu	Val	Ser	Arg	Tyr	Pro	Asp	Leu	Ser	Leu	Val	
				3125					3130					3135		
Lys	Glu	Glu	Pro	Pro	Glu	Pro	Val	Pro	Ser	Pro	Ile	Ile	Pro	Ile	Leu	
			3140					3145					3150			
Pro	Ser	Thr	Ala	Gly	Lys	Ser	Ser	Glu	Ser	Arg	Arg	Asn	Asp	Ile	Lys	
		3155				3160					3165					
Thr	Glu	Pro	Gly	Thr	Leu	Tyr	Phe	Ala	Ser	Pro	Phe	Gly	Pro	Ser	Pro	
		3170				3175					3180					
Asn	Gly	Pro	Arg	Ser	Gly	Leu	Ile	Ser	Val	Ala	Ile	Thr	Leu	His	Pro	

3185 3190 3195 3200
 Thr Ala Ala Glu Asn Ile Ser Ser Val Val Ala Ala Phe Ser Asp Leu
 3205 3210 3215
 Leu His Val Arg Ile Pro Asn Ser Tyr Glu Val Ser Ser Ala Pro Asp
 3220 3225 3230
 Val Pro Ser Met Gly Leu Val Ser Ser His Arg Ile Asn Pro Gly Leu
 3235 3240 3245
 Glu Tyr Arg Gln His Leu Leu Leu Arg Gly Pro Pro Pro Gly Ser Ala
 3250 3255 3260
 Asn Pro Pro Arg Leu Val Ser Ser Tyr Arg Leu Lys Gln Pro Asn Val
 3265 3270 3275 3280
 Pro Phe Pro Pro Thr Ser Asn Gly Leu Ser Gly Tyr Lys Asp Ser Ser
 3285 3290 3295
 His Gly Ile Ala Glu Ser Ala Ala Leu Arg Pro Gln Trp Cys Cys His
 3300 3305 3310
 Cys Lys Val Val Ile Leu Gly Ser Gly Val Arg Lys Ser Phe Lys Asp
 3315 3320 3325
 Leu Thr Leu Leu Asn Lys Asp Ser Arg Glu Ser Thr Lys Arg Val Glu
 3330 3335 3340
 Lys Asp Ile Val Phe Cys Ser Asn Asn Cys Phe Ile Leu Tyr Ser Ser
 3345 3350 3355 3360
 Thr Ala Gln Ala Lys Asn Ser Glu Asn Lys Glu Ser Ile Pro Ser Leu
 3365 3370 3375
 Pro Gln Ser Pro Met Arg Glu Thr Pro Ser Lys Ala Phe His Gln Tyr
 3380 3385 3390
 Ser Asn Asn Ile Ser Thr Leu Asp Val His Cys Leu Pro Gln Leu Pro
 3395 3400 3405
 Glu Lys Ala Ser Pro Pro Ala Ser Pro Pro Ile Ala Phe Pro Pro Ala
 3410 3415 3420
 Phe Glu Ala Ala Gln Val Glu Ala Lys Pro Asp Glu Leu Lys Val Thr
 3425 3430 3435 3440
 Val Lys Leu Lys Pro Arg Leu Arg Ala Val His Gly Gly Phe Glu Asp
 3445 3450 3455
 Cys Arg Pro Leu Asn Lys Lys Trp Arg Gly Met Lys Trp Lys Lys Trp
 3460 3465 3470
 Ser Ile His Ile Val Ile Pro Lys Gly Thr Phe Lys Pro Pro Cys Glu

000000"000000"000000

3475					3480					3485						
Asp	Glu	Ile	Asp	Glu	Phe	Leu	Lys	Lys	Leu	Gly	Thr	Ser	Leu	Lys	Pro	
3490					3495					3500						
Asp	Pro	Val	Pro	Lys	Asp	Tyr	Arg	Lys	Cys	Cys	Phe	Cys	His	Glu	Glu	
3505					3510					3515					3520	
Gly	Asp	Gly	Leu	Thr	Asp	Gly	Pro	Ala	Arg	Leu	Leu	Asn	Leu	Asp	Leu	
					3525					3530					3535	
Asp	Leu	Trp	Val	His	Leu	Asn	Cys	Ala	Leu	Trp	Ser	Thr	Glu	Val	Tyr	
					3540					3545					3550	
Glu	Thr	Gln	Ala	Gly	Ala	Leu	Ile	Asn	Val	Glu	Leu	Ala	Leu	Arg	Arg	
					3555					3560					3565	
Gly	Leu	Gln	Met	Lys	Cys	Val	Phe	Cys	His	Lys	Thr	Gly	Ala	Thr	Ser	
					3570					3575					3580	
Gly	Cys	His	Arg	Phe	Arg	Cys	Thr	Asn	Ile	Tyr	His	Phe	Thr	Cys	Ala	
					3585					3590					3595	
Ile	Lys	Ala	Gln	Cys	Met	Phe	Phe	Lys	Asp	Lys	Thr	Met	Leu	Cys	Pro	
					3605					3610					3615	
Met	His	Lys	Pro	Lys	Gly	Ile	His	Glu	Gln	Glu	Leu	Ser	Tyr	Phe	Ala	
					3620					3625					3630	
Val	Phe	Arg	Arg	Val	Tyr	Val	Gln	Arg	Asp	Glu	Val	Arg	Gln	Ile	Ala	
					3635					3640					3645	
Ser	Ile	Val	Gln	Arg	Gly	Glu	Arg	Asp	His	Thr	Phe	Arg	Val	Gly	Ser	
					3650					3655					3660	
Leu	Ile	Phe	His	Thr	Ile	Gly	Gln	Leu	Leu	Pro	Gln	Gln	Met	Gln	Ala	
					3665					3670					3675	
Phe	His	Ser	Pro	Lys	Ala	Leu	Phe	Pro	Val	Gly	Tyr	Glu	Ala	Ser	Arg	
					3685					3690					3695	
Leu	Tyr	Trp	Ser	Thr	Arg	Tyr	Ala	Asn	Arg	Arg	Cys	Arg	Tyr	Leu	Cys	
					3700					3705					3710	
Ser	Ile	Glu	Glu	Lys	Asp	Gly	Arg	Pro	Val	Phe	Val	Ile	Arg	Ile	Val	
					3715					3720					3725	
Glu	Gln	Gly	His	Glu	Asp	Leu	Val	Leu	Ser	Asp	Ile	Ser	Pro	Lys	Gly	
					3730					3735					3740	
Val	Trp	Asp	Lys	Ile	Leu	Glu	Pro	Val	Ala	Cys	Val	Arg	Lys	Lys	Ser	
					3745					3750					3755	
Glu	Met	Leu	Gln	Leu	Phe	Pro	Ala	Tyr	Leu	Lys	Gly	Glu	Asp	Leu	Phe	

<210>	426
<211>	174
<212>	PRT

<400> 426

<400> 427

Pro Pro Ala Gln Leu Tyr His Trp Val Glu Met Arg Thr Lys Met Arg
50 55 60

Ile Met Gly Phe Arg Gly Thr Val Ile Lys Pro Leu Asn Glu Glu Ala
 65 70 75 80
 Ala Ala Glu Leu Gly Ala Glu Leu Leu Gly Glu Ala Thr Ile Phe Ile
 85 90 95
 Val Gly Gly Gly Cys Leu Val Leu Glu Tyr Trp Arg His Gln Ala Gln
 100 105 110
 Gln Arg His Lys Glu Glu Glu Gln Arg Ala Ala Trp Asn Ala Leu Arg
 115 120 125
 Asp Glu Val Gly His Leu Ala Leu Ala Leu Glu Ala Leu Gln Ala Gln
 130 135 140
 Val Gln Ala Ala Pro Pro Gln Gly Ala Leu Glu Glu Leu Arg Thr Glu
 145 150 155 160
 Leu Gln Glu Val Arg Ala Gln Leu Cys Asn Pro Gly Arg Ser Ala Ser
 165 170 175
 His Ala Val Pro Ala Ser Lys Lys
 180

<210> 428
 <211> 6476
 <212> DNA
 <213> Homo sapiens

<400> 428
 cactgactgg actgaaaaca gggccaagaa aactgctgct gcaggggggtc ctgaaaacag 60
 ctggaacccg gcagtgatgt gggacctaac ttgaagttaa cctgtggtgg tgaggttgga 120
 accagttgga ttatgattta ttttctacac tctgtacgg aatgcagagc tgttgatcc 180
 tgatgaatct actgctaaat atagtcattt ggaataattt taagtattga tcttaaaact 240
 tgtaccacaa caagagtgtc taaaaagcac ggcaagctca ttacgttctt acgaacattc 300
 atgaagtctc gtccaacaaa acagaagctg aagcagcggg gaatcttgaa agagaggggtg 360
 tttggttggtg acctggggga gcaccttcta aattctgggt ttgaagtgcc gcagttcttc 420
 aaagctgcac agcattcatt gagagatatg gcacgttgga tggaatctat cgcctttctg 480
 gtgttgccctc caatatccag agactacgcc atgaatttga ctctgagcac gtccccgacc 540
 tgacgaaaga accgtatgtt caggacatcc attctgtggg ttccctatgt aagctgtact 600
 tccgggaact cccaaaccct ctgcttacct accagctgta tgagaaattt tctgatgcag 660
 tttcagcagc aacagatgaa gaaaggctga taaaaatcca cgatgtcatc cagcagctcc 720
 cccaccaca ctacagaaca ctggagttcc tgatgagaca cttgtctctt ctactgtact 780
 attgttccat cacaatatg catgcaaaaa atctagcaat tgtttgggct ccaaacctgt 840
 taagatcaaa acagatagaa tctgcctgct tcagtggaa acagcgtttc atggaagtga 900
 ggattcagtc tgtggttggt gagttcatcc tgaatcacgt tgatgtgctg ttcagcggca 960
 gaatcagcat ggccatgcaa gagggggcag cttctctatc aaggcccaag tccctcctgg 1020
 tctcctctcc atccacaaa ctgctgacat tggaagaggg ccaggcacga acacaagctc 1080

aggtcaattc	tccaattgtg	acggaataa	aatatatcga	agtaggagaa	ggacctgctg	1140
cacttcaggg	gaaatttcct	accataattg	agttcccact	tgaagagaa	aggcctcaaa	1200
ataagatgaa	aaagtctcct	gtgggtagct	ggcgttcctt	tttcaacttg	gggaaatcat	1260
catctgtttc	taaacgaaag	ctgcagcgga	atgagagtga	gccttcagag	atgaaagcca	1320
tggctctgaa	aggtggcagg	gcagaaggaa	ccctccgttc	agctaaaagt	gaggagtctc	1380
ttacatctct	ccatgcagtt	gatggtgatt	ctaagctctt	ccgaccaga	agaccagat	1440
cttcacagtga	tgcactgtct	gcctctttta	atggagaaat	gctggggaac	cgctgtaact	1500
cctatgataa	tctgcctcat	gacaatgaga	gtgaggagga	aggagggctg	cttcatatcc	1560
cagcccttat	gtctcctcat	tcagctgagg	atggtgactt	gagccacca	gacattggag	1620
tagccagcct	ggattttgat	ccaatgtcat	ttcaatgtag	tcctcctaag	gccgaatcag	1680
aatgtctgga	gagtgggtgt	tccttttttag	attcaccagg	atactccaag	gataaaccaa	1740
gtgccaataa	aaaggatgca	gaaacaggta	gtagccaatg	tcagactcca	ggaagcacag	1800
caagctctga	acctgtctct	cctcttcagg	agaaaactgag	tccattcttt	accctggact	1860
tgagcccaac	tgaagataaa	tcactaagc	catectcctt	tactgaaaag	gtcgtctatg	1920
ctttctctcc	gaagatagga	cggaaattaa	gcaaatacc	ttctatgagc	atatctgagc	1980
caatttcagt	gaccctacca	ccacgggtgt	cagaagtcat	tggtacagt	tcaaatacca	2040
cagctcagaa	tgcactcatc	tcaacctggg	acaaatgcgt	tgaagaaaag	gatgccacaa	2100
atagatcccc	caccacagata	gtaaaagta	aaaaaactga	gacagttgcc	caagaagcat	2160
atgaatctga	agtcagagccc	ctggaccagg	tggtgctgta	agaagtagaa	ttgccaggga	2220
aagaggatca	gtctgtctca	agcagtcaga	gtaaggctgt	agcttctgga	cagactcaga	2280
caggagcagt	taccatgac	ccccctcagg	attccgttcc	tgtcagttca	gtctctctta	2340
tcccaccacc	accgctccg	aaaaatgttg	ccggaatggt	ggcgctagca	ttagctgagt	2400
ccgcacagca	agcctcaact	cagtcattga	agagaccagg	gacctctcag	gctgggtata	2460
caaattatgg	agacatagcg	gtggctacaa	ctgaagataa	tctgtccagt	tcttactctg	2520
cagttgctct	agataaggcc	tatttccaaa	ccgatcgacc	agcagagcag	ttccacctcc	2580
agaataatgc	accaggaaac	tgtgaccatc	ctctaccaga	gacaacagct	actggggatc	2640
ctaccattc	caacacaact	gaatctgggg	agcaacatca	ccaagtagac	ttaacaggga	2700
atcagccaca	tcaagcatat	ttatctgggg	accagaaaa	ggccagaatt	acttcagttc	2760
ccttagactc	agagaagtct	gatgatcatg	taagtttccc	tgaagaccag	tctgggaaga	2820
acagtatgcc	aactgtctcc	ttcttggtgc	aggaccagtc	tccacccctg	ttctacagtg	2880
gagatcagcc	tccttcttat	cttggtgcaa	gtgtggataa	actccatcac	cctttagaat	2940
ttgcagacaa	atctcccaca	cctctaatt	tacctagcga	taaaatctac	cctccttctg	3000
ggtcccccca	agagaatacc	agcacagcca	ccatgactta	catgacaact	actccagcaa	3060
cagcccaaat	gagcaccagc	gaagccagct	gggagtgggc	tgaacaaccc	accactgctg	3120
atthtctgct	tgccacactt	cagcgcagc	acagaaacta	tcgtcccctt	ccccctccgc	3180
cttcccagag	atctgcagag	cagccaccag	ttgtggggca	ggtacaagca	gcaaccaata	3240
taggattaaa	taattcccac	aaggttcaag	gagtagttcc	agttccagag	aggccacctg	3300
aacctcgagc	catggatgac	cctgcgtctg	ccttcatcag	tgacagtggg	gctgctgctg	3360
ctcagtgctc	catggctaca	gctgtccagc	caggcctgcc	tgagaaagtg	cgggacgggtg	3420
cccgggtccc	gctgctgcac	ctgcgcgccg	agtctgtccc	tgcgcatecc	tgtggctttc	3480
ctgcaccact	gccccccacc	aggatgatgg	agagtaagat	gattgctgcc	ataactcca	3540
gcagtgcaga	tgccaccagc	agttcaaatt	atcatctcct	tgtcactgct	tcattccact	3600
ctgtggaaga	tgcattgcct	ttaccacttc	ctgtcccaca	acctaaagcat	gcttctcaga	3660
aaacagttta	ctcctccttt	gctaggcccg	atgtcaccac	tgaacctctt	ggtccagata	3720
actgttttga	tttcaatatg	actccaaact	gccagtaccg	tcccagagt	gtacctcccc	3780
atcacaataa	attggagcag	caccaagtgt	atggtgccag	gtcagagcca	ccagcctcca	3840
tgggtcttcg	ttataacaca	tatgtggccc	caggaagaaa	cgcatctgga	caccactcca	3900
agccatgcag	ccgggtcgag	tatgtgtctt	ctttgagctc	ctctgtcagg	aataacctgtt	3960
accccgaaga	cattccaccg	tacctacca	tccggagagt	gcagttcttc	catgctccgc	4020
cgtcttccat	gattcgtctc	gttcccattt	cacggacaga	agttccccca	gatgatgagc	4080
cagcctactg	cccaagacct	ctgtaccaat	ataagccata	tcagtcctcc	caggcccgct	4140
cagattatca	tgtcactcag	cttcagcctt	actttgagaa	tggccgggtc	cactacaggt	4200
atagcccata	ttccagttct	tctagtctct	attacagtcc	agatggggcc	ctgtgtgatg	4260
tggatgccta	tggcacagtc	cagttgagac	cccttcaccg	ccttcccaat	cgagactttg	4320

cggccacctg gatccctgcc cctccccact gggacggaat aaatgctctg cagacctgaa 720
 aaaaaaaaaa aa 732

<210> 430
 <211> 2843
 <212> DNA
 <213> Homo sapiens

<400> 430
 gcgggcccag gaggcggcgg cgggcgcggc ggacgggccc cccgcggcag acggcgagga 60
 cggacaggac ccgcacagca agcacctgta cacggccgac atgttcacgc acgggatcca 120
 gagcgcgcgc cacttcgtca tgttcttcgc gccctggtgt ggacactgcc agcggtgca 180
 gccgacttgg aatgacctgg gagacaaata caacagcatg gaagatgcca aagtctatgt 240
 ggctaaagtg gactgcacgg cccactccga cgtgtgctcc gccagggggg tgcgaggata 300
 ccccacctta aagcttttca agccaggcca agaagctgtg aagtaccagg gtccctcgga 360
 cttccagaca ctggaaaact ggatgctgca gacactgaac gaggagccag tgacaccaga 420
 gccggaagtg gaaccgcccc gtgcccccca gctcaagcaa gggctgtatg agctctcagc 480
 aagcaacttt gagctgcacg ttgcacaagg cgaccacttt atcaagttct tcgctccgtg 540
 gtgtggtcac tgcaaagccc tggctccaac ctgggagcag ctggctctgg gccttgaaca 600
 ttccgaaaact gtcaagattg gcaaggttga ttgtacacag cactatgaac tctgctccgg 660
 aaaccagggt ctgtggtatc ccactcttct ctggttccga gatgggaaaa aggtggatca 720
 gtacaaggga aagcgggatt tggagtcact gagggagtac gtggagtgc agctgcagcg 780
 cacagagact ggagcgacgg agaccgtcac gccctcagag gccccggtgc tggcagctga 840
 gcccgaggct gacaagggca ctgtgttggc actcaatgaa aatactttcg atgacaccat 900
 tgcagaagga ataaccttca tcaagtttta tgcctccatgg tgtggtcatt gtaagactct 960
 ggctcctact tgggagggaac tctctaaaaa ggaattccct ggtctggcgg ggtcaagat 1020
 cgccgaagta gactgcactg ctgaacggaa tatctgcagc aagtattcgg tacgaggcta 1080
 cccacgtta ttgcttttcc gaggagggaa gaaagtcagt gagcacagtg gaggcagaga 1140
 ccttgactcg ttacaccgct ttgtcctgag ccaagcgaaa gacgaacttt aggaacacag 1200
 ttggaggtca cctctcctgc ccagctcccg caccctgcgt ttaggagttc agtcccacag 1260
 aggccactgg gttcccagtg gtggctgttc agaaagcaga acatactaag cgtgaggtat 1320
 cttctttgtg tgtgtgtttt ccaagccaac acactctaca gattctttat taagttaagt 1380
 ttctctaagt aaatgtgtaa ctcatggtca ctgtgtaaac attttcagtg gcgatatatc 1440
 ccctttgacc ttctcttgat gaaatttaca tggtttcctt tgagactaaa atagcgttga 1500
 gggaaatgaa attgctggac tatttgtggc tcctgagttg agtgattttg gtgaaagaaa 1560
 gcacatccaa agcatagttt acctgcccac gagttctgga aaggtggcct tgtggcagta 1620
 ttgacgttcc tctgatctta aggtcacagt tgactcaata ctgtgttggc ccgtagcatg 1680
 gagcagattg aaatgcaaaa acccacacct ctggaagata cttcacggc cgctgctgga 1740
 gcttctgttg ctgtgaatac ttctctcagt gtgagaggtt agccgtgatg aaagcagcgt 1800
 tacttctgac cgtgcctgag taagagaatg ctgatgccat aactttatgt gtcgatactt 1860
 gtcaaatcag ttactgttca ggggatcctt ctgtttctca cggggtgaaa catgtcttta 1920
 gttcctcatg ttaacacgaa gccagagccc acatgaactg ttggatgtct tccttagaaa 1980
 gggtaggcat ggaaaaattcc acgaggctca ttctcagtat ctcatctaact cattgaaaga 2040
 ttccagttgt atttgtcacc tgggggtgaca agaccagaca ggctttccca ggctgggta 2100
 tccagggagg ctctgcagcc ctgctgaagg gccctaacta gagttctaga gtttctgatt 2160
 ctgtttctca gtagtcttt tagaggcttg ctatacttgg tctgcttcaa ggaggtcgac 2220
 cttctaattgt atgaagaatg ggatgcattt gatctcaaga ccaaagacag atgtcagtgg 2280
 gctgctctgg ccctgggtgtg cacggctgtg gcagctgttg atgccagtgt cctctaactc 2340
 atgctgtcct tgtgattaaa cacctctatc tcccttggga ataagcacat acaggcttaa 2400
 gctctaagat agataggtgt ttgtcctttt accatcgagc tacttcccat aataaccact 2460
 ttgcatccaa cactcttcac ccacctccca tacgcaaggg gatgtggata cttggcccaa 2520
 agtaactggt ggtaggaatc ttagaaacaa gaccacttat actgtctgtc tgaggcagaa 2580
 gataacagca gcatctcgac cagcctctgc cttaaaggaa atctttatta atcacgtatg 2640
 gttcacagat aattcttttt ttaaaaaaac ccaacctcct agagaagcac aactgtcaag 2700

```

agtcttgtac acacaacttc agctttgcat cactgagctt gtattccaag aaaatcaaaag 2760
tggtacaatt tgtttgttta cactatgata ctttctaaat aaactctttt ttttaaaaaa 2820
aaaaaaaaa aaaaaaactc gag                                     2843

```

```

<210> 431
<211> 640
<212> DNA
<213> Homo sapiens

```

```

<400> 431
ggtaacgtta tagtatttgt cagaagtttg ggtctccgtg ggcattgtga tccgtcccag 60
gcagtggatt aggaggccag aaggagatcc cttccacggt gctaggctga gatggatcct 120
ctcagggccc aacagctggc tgcggagctg gaggtggaga tgatggccga tatgtacaac 180
agaatgacca gtgcctgcca ccggaagtgt gtgcctcctc actacaagga agcagagctc 240
tccaagggcg agtctgtgtg cctggaccga tgtgtctcta agtacctgga catccatgag 300
cggatgggca aaaagttgac agagttgtct atgcaggatg aagagctgat gaagaggggtg 360
cagcagagct ctgggctgtc atgaggtccc tgtcagtata caccctgggg tgtacccac 420
cccttccac ttttaataaac gtgctccctg ttgggtgtca tctgtgaaga ctgccaggcc 480
taggctctct gtagagagtc ttcaagatcc cggagtggta gcgctgtctc ctggtgaagg 540
agtatttgtc aacttggat gtgactgtgt gtgtatgtat gtgtatatat atatatatat 600
atatatataa acaagtttgt tgacacctac aaaaaaaaaa 640

```

```

<210> 432
<211> 2068
<212> DNA
<213> Homo sapiens

```

```

<400> 432
cctcagaagt ccgtgccagt gaccggaggg ggccggggcg agcggttcct tgtgggctag 60
aagaatcctg caaaaatgtc tctctatcca tctctcgaag acttgaaggt agacaaagta 120
attcaggctc aaactgcttt ttctgcaaac cctgccaatc cagcaatttt gtcagaagct 180
tctgtctcta tccctcacga tggaaatctc tatcccagac tgtatccaga gctctctcaa 240
tacatggggc tgagttaaaa tgaagaagaa atacgtgcaa atgtggccgt ggtttctggt 300
gcaccacttc aggggcagtt ggtagcaaga ccttccagta taaactatat ggtggctcct 360
gtaactggta atgatgttgg aattcgtaga gcagaaatta agcaagggat tctgtgaagtc 420
attttgtgta aggatcaaga tggaaaaatt ggactcaggc ttaaataaat agataatggt 480
atatttgttc agctagtcca ggctaattct ccagcctcat tggttggctc gagatttggg 540
gaccaagtac ttcagatcaa tggtgaaaac tgtgcaggat ggagctctga taaagcgcac 600
aaggtgctca aacaggcttt tggagagaag attaccatga ccattcgtga caggcccttt 660
gaacggacga ttaccatgca taaggatagc actggacatg ttggttttat ctttaaaaaa 720
ggaaaaataa catccatagt gaaagatagc tctgcagcca gaaatggctc tctcacggaa 780
cataacatct gtgaaatcaa tggacagaat gtcattggat tgaaggactc tcaaattgca 840
gacatactgt caacatctgg gactgtagtt actattacaa tcatgcctgc ttttatcttt 900
gaacatatta ttaagcggat ggcaccaagc attatgaaaa gcctaattgga ccacaccatt 960
cctgaggttt aaaattcacg gcacctgga aatgtagctg aacgtctcca gtttccttct 1020
ttggcaactt ctgtattatg cacgtgaagc cttcccgagg ccagcgagca tatgctgcat 1080
gaggaccttt ctatcttaca ttatggctgg gaatcttact ctttcatctg ataccttggt 1140
cagatttcaa aatagttgta gccttatcct ggttttacag atgtgaaact ttcaagagat 1200
ttactgactt tctagaata gtttctctac tggaaacctg atgcttttat aagccattgt 1260
gattaggatg actgttacag gcttagcttt gtgtgaaaac cagtcacctt tctcctaggt 1320
aatgagtagt gctgttcata ttactttagt tctatagcat actgcatctt taacatgcta 1380
tcatagtaca tttagaatga ttgcctttga tttttttttt aaattctgtg tgtgtgtgtg 1440
taaaatgcc aattaagaaca ctggtttcat tccatgtaag cattaacacag tgtatgtagg 1500
tttcaagaga ttgtgatgat tcttaaattt taactacctt caottaatat gcttgaactg 1560

```

```

tcgccttaac tatgttaagc atctagacta aaagccaaaa tataattatt gctgcctttc 1620
taaaaaacca aaatgtagtt ctctattaac ctgaaatgta cactagccca gaacagttta 1680
atggtactta ctgagctata gcatagctgc ttagttgttt ttgagagttt ttagtcaaca 1740
cataatggaa acttctttct tctaaaagtt gccagtgcc a ttttaagaa gtgaatcact 1800
atatgtgatg taaaagttat tacactaaac aggataaaact tttgactccc cttttgttca 1860
tttgtggatt aagtgggtata atacttaatt ttggcatttg actcttaaga ttatgtaacc 1920
tagctacttt gggatgggtct tagaatattt ttctgataac ttgttccttt tcctgactcc 1980
tccttgcaaa caaaatgata gttgacactt tatcctgatt tttttcttct ttttggttta 2040
tgtctattct aattaaatat gtataaat 2068

```

```

<210> 433
<211> 1723
<212> DNA
<213> Homo sapiens

```

```

<400> 433
tttctttgtt aagtcgttcc ctctacaaag gacttcctag tgggtgtgaa aggcagcggg 60
ggccacagag gcggcggaga gatggccttc agcggttccc aggctcccta cctgagtcca 120
gctgtccctt tttctgggac tattcaagga ggtctccagg acggacttca gatcactgtc 180
aatgggaccg ttctcagctc cagtggaaac aggtttgctg tgaactttca gactggcttc 240
agtggaaatg acattgcctt ccacttcaac cctcggtttg aagatggagg gtacgtggtg 300
tgcaacacga ggcagaacgg aagctggggg cccgaggaga ggaagacaca catgcctttc 360
cagaagggga tgccctttga cctctgcttc ctggtgcaga gctcagattt caaggtgatg 420
gtgaacggga tcctcttcgt gcagttactc caccgcgtgc ccttccaccg tgtggacacc 480
atctccgtca atggctctgt gcagctgtcc tacatcagct tccagaaccc ccgcacagtc 540
cctgttcagc ctgcctttcc acggtgcccgt tctcccagcc tgtctgtttc ccaccaggc 600
ccagggggcg cagacaaaaa cctcccggcg tgtggcctgc caaccgggt cccattaccc 660
agacagtcat ccacacagtg cagagcgccc ctggacagat gttctctact cccgccatcc 720
cacctatgat gtacccccac ccgcctatc cgatgccttt catcaccacc attctgggag 780
ggctgtaccc atccaagtc atcctcctgt caggcactgt cctgccaggt gctcagaggt 840
tccacatcaa cctgtgctct gggaaccaca tcgccttcca cctgaacccc cgttttgatg 900
agaatgctgt ggtccgcaac acccagatcg acaactcctg ggggtctgag gagcgaagtc 960
tgccccgaaa aatgcccttc gtccgtggcc agagcttctc agtgtggatc ttgtgtgaag 1020
ctcactgcct caaggtggcc gtggatggtc agcacctgtt tgaatactac catcgctga 1080
ggaacctgcc caccatcaac agactggaag tggggggcga catccagctg acccatgtgc 1140
agacataggc ggcttcctgg ccctggggcc gggggctggg gtgtggggca gtctgggtcc 1200
tctcatcatc cccacttccc aggccagcc tttccaaccc tgccctgggat ctgggcttta 1260
atgcagaggc catgtccttg tctggctcctg cttctggcta cagccacctt ggaacggaga 1320
aggcagctga cggggattgc ctctctcagc cgcagcagca cctggggctc cagctgctgg 1380
aatcctacca tcccaggagg caggcacagc caggagagg ggaggagtgg gcagtgaaga 1440
tgaagcccca tgctcagtcc cctcccatcc cccacgcagc tccacccag tcccaagcca 1500
ccagctgtct gctcctgggt ggaggtggcc tctcagccc ctctctctg accttaacc 1560
tcactctcac cttgcaccgt gcaccaaccc ttcacccctc ctggaaagca ggctgatgg 1620
cttcccactg gcctccacca cctgaccaga gtgttctctt cagaggactg gctcctttcc 1680
cagtgtcctt aaaataaaga aatgaaaatg cttgttgga cat 1723

```

```

<210> 434
<211> 1702
<212> PRT
<213> Homo sapiens

```

```

<400> 434
Ala Ala Val Leu Gln Ser Cys Thr Ala Phe Ile Glu Arg Tyr Gly Ile
      5              10              15

```


Asp Gln Val Ala Ala Glu Glu Val Glu Leu Pro Gly Lys Glu Asp Gln
 595 600 605
 Ser Val Ser Ser Ser Gln Ser Lys Ala Val Ala Ser Gly Gln Thr Gln
 610 615 620
 Thr Gly Ala Val Thr His Asp Pro Pro Gln Asp Ser Val Pro Val Ser
 625 630 635 640
 Ser Val Ser Leu Ile Pro Pro Pro Pro Pro Lys Asn Val Ala Arg
 645 650 655
 Met Leu Ala Leu Ala Leu Ala Glu Ser Ala Gln Gln Ala Ser Thr Gln
 660 665 670
 Ser Leu Lys Arg Pro Gly Thr Ser Gln Ala Gly Tyr Thr Asn Tyr Gly
 675 680 685
 Asp Ile Ala Val Ala Thr Thr Glu Asp Asn Leu Ser Ser Ser Tyr Ser
 690 695 700
 Ala Val Ala Leu Asp Lys Ala Tyr Phe Gln Thr Asp Arg Pro Ala Glu
 705 710 715 720
 Gln Phe His Leu Gln Asn Asn Ala Pro Gly Asn Cys Asp His Pro Leu
 725 730 735
 Pro Glu Thr Thr Ala Thr Gly Asp Pro Thr His Ser Asn Thr Thr Glu
 740 745 750
 Ser Gly Glu Gln His His Gln Val Asp Leu Thr Gly Asn Gln Pro His
 755 760 765
 Gln Ala Tyr Leu Ser Gly Asp Pro Glu Lys Ala Arg Ile Thr Ser Val
 770 775 780
 Pro Leu Asp Ser Glu Lys Ser Asp Asp His Val Ser Phe Pro Glu Asp
 785 790 795 800
 Gln Ser Gly Lys Asn Ser Met Pro Thr Val Ser Phe Leu Asp Gln Asp
 805 810 815
 Gln Ser Pro Pro Arg Phe Tyr Ser Gly Asp Gln Pro Pro Ser Tyr Leu
 820 825 830
 Gly Ala Ser Val Asp Lys Leu His His Pro Leu Glu Phe Ala Asp Lys
 835 840 845
 Ser Pro Thr Pro Pro Asn Leu Pro Ser Asp Lys Ile Tyr Pro Pro Ser
 850 855 860
 Gly Ser Pro Glu Glu Asn Thr Ser Thr Ala Thr Met Thr Tyr Met Thr
 865 870 875 880

000250"022950

Arg	Val	Glu	Tyr	Val	Ser	Ser	Leu	Ser	Ser	Ser	Val	Arg	Asn	Thr	Cys		
1170					1175					1180							
Tyr	Pro	Glu	Asp	Ile	Pro	Pro	Tyr	Pro	Thr	Ile	Arg	Arg	Val	Gln	Ser		
1185					1190					1195						1200	
Leu	His	Ala	Pro	Pro	Ser	Ser	Met	Ile	Arg	Ser	Val	Pro	Ile	Ser	Arg		
1205					1210					1215							
Thr	Glu	Val	Pro	Pro	Asp	Asp	Glu	Pro	Ala	Tyr	Cys	Pro	Arg	Pro	Leu		
1220					1225					1230							
Tyr	Gln	Tyr	Lys	Pro	Tyr	Gln	Ser	Ser	Gln	Ala	Arg	Ser	Asp	Tyr	His		
1235					1240					1245							
Val	Thr	Gln	Leu	Gln	Pro	Tyr	Phe	Glu	Asn	Gly	Arg	Val	His	Tyr	Arg		
1250					1255					1260							
Tyr	Ser	Pro	Tyr	Ser	Ser	Ser	Ser	Ser	Ser	Tyr	Tyr	Ser	Pro	Asp	Gly		
1265					1270					1275						1280	
Ala	Leu	Cys	Asp	Val	Asp	Ala	Tyr	Gly	Thr	Val	Gln	Leu	Arg	Pro	Leu		
1285					1290					1295							
His	Arg	Leu	Pro	Asn	Arg	Asp	Phe	Ala	Phe	Tyr	Asn	Pro	Arg	Leu	Gln		
1300					1305					1310							
Gly	Lys	Ser	Leu	Tyr	Ser	Tyr	Ala	Gly	Leu	Ala	Pro	Arg	Pro	Arg	Ala		
1315					1320					1325							
Asn	Val	Thr	Gly	Tyr	Phe	Ser	Pro	Asn	Asp	His	Asn	Val	Val	Ser	Met		
1330					1335					1340							
Pro	Pro	Ala	Ala	Asp	Val	Lys	His	Thr	Tyr	Thr	Ser	Trp	Asp	Leu	Glu		
1345					1350					1355						1360	
Asp	Met	Glu	Lys	Tyr	Arg	Met	Gln	Ser	Ile	Arg	Arg	Glu	Ser	Arg	Ala		
1365					1370					1375							
Arg	Gln	Lys	Val	Lys	Gly	Pro	Val	Met	Ser	Gln	Tyr	Asp	Asn	Met	Thr		
1380					1385					1390							
Pro	Ala	Val	Gln	Asp	Asp	Leu	Gly	Gly	Ile	Tyr	Val	Ile	His	Leu	Arg		
1395					1400					1405							
Ser	Lys	Ser	Asp	Pro	Gly	Lys	Thr	Gly	Leu	Leu	Ser	Val	Ala	Glu	Gly		
1410					1415					1420							
Lys	Glu	Ser	Arg	His	Ala	Ala	Lys	Ala	Ile	Ser	Pro	Glu	Gly	Glu	Asp		
1425					1430					1435						1440	
Arg	Phe	Tyr	Arg	Arg	His	Pro	Glu	Ala	Glu	Met	Asp	Arg	Ala	His	His		
1445					1450					1455							

His Gly Gly His Gly Ser Thr Gln Pro Glu Lys Pro Ser Leu Pro Gln
 1460 1465 1470
 Lys Gln Ser Ser Leu Arg Ser Arg Lys Leu Pro Asp Met Gly Cys Ser
 1475 1480 1485
 Leu Pro Glu His Arg Ala His Gln Glu Ala Ser His Arg Gln Phe Cys
 1490 1495 1500
 Glu Ser Lys Asn Gly Pro Pro Tyr Pro Gln Gly Ala Gly Gln Leu Asp
 1505 1510 1515 1520
 Tyr Gly Ser Lys Gly Ile Pro Asp Thr Ser Glu Pro Val Ser Tyr His
 1525 1530 1535
 Asn Ser Gly Val Lys Tyr Ala Ala Ser Gly Gln Glu Ser Leu Arg Leu
 1540 1545 1550
 Asn His Lys Glu Val Arg Leu Ser Lys Glu Met Glu Arg Pro Trp Val
 1555 1560 1565
 Arg Gln Pro Ser Ala Pro Glu Lys His Ser Arg Asp Cys Tyr Lys Glu
 1570 1575 1580
 Glu Glu His Leu Thr Gln Ser Ile Val Pro Pro Pro Lys Pro Glu Arg
 1585 1590 1595 1600
 Ser His Ser Leu Lys Leu His His Thr Gln Asn Val Glu Arg Asp Pro
 1605 1610 1615
 Ser Val Leu Tyr Gln Tyr Gln Pro His Gly Lys Arg Gln Ser Ser Val
 1620 1625 1630
 Thr Val Val Ser Gln Tyr Asp Asn Leu Glu Asp Tyr His Ser Leu Pro
 1635 1640 1645
 Gln His Gln Arg Gly Val Phe Gly Gly Gly Gly Met Gly Thr Tyr Val
 1650 1655 1660
 Pro Pro Gly Phe Pro His Pro Gln Ser Arg Thr Tyr Ala Thr Ala Leu
 1665 1670 1675 1680
 Gly Gln Gly Ala Phe Leu Pro Ala Glu Leu Ser Leu Gln His Pro Glu
 1685 1690 1695
 Thr Gln Ile His Ala Glu
 1700

<210> 435
 <211> 160
 <212> PRT
 <213> Homo sapiens

000260"07"0960

Pro Phe Gln Gln Val Gly Arg Cys Asn Pro Ser Pro Gln Thr Arg Pro
5 10 15

Gly Pro Ala Ser Lys Val Lys Gln Asp Met Pro Pro Pro Gly Gly Tyr
20 25 30

Gly Pro Ile Asp Tyr Lys Arg Asn Leu Pro Arg Arg Gly Leu Ser Gly
35 40 45

Tyr Ser Met Leu Ala Ile Gly Ile Gly Thr Leu Ile Tyr Gly His Trp
50 55 60

Ser Ile Met Lys Trp Asn Arg Glu Arg Arg Arg Leu Gln Ile Glu Asp
65 70 75 80

Phe Glu Ala Arg Ile Ala Leu Leu Pro Leu Leu Gln Ala Glu Thr Asp
85 90 95

Arg Arg Thr Leu Gln Met Leu Arg Glu Asn Leu Glu Glu Glu Ala Ile
100 105 110

Ile Met Lys Asp Val Pro Asp Trp Lys Val Gly Glu Ser Val Phe His
115 120 125

Thr Thr Arg Trp Val Pro Pro Leu Ile Gly Glu Leu Tyr Gly Leu Arg
130 135 140

Thr Thr Glu Glu Ala Leu His Ala Ser His Gly Phe Met Trp Tyr Thr
145 150 155 160

<211> 396

<212> PRT

<213> Homo sapiens

Arg Ala Gln Glu Ala Ala Ala Ala Ala Asp Gly Pro Pro Ala Ala
5 10 15

Asp Gly Glu Asp Gly Gln Asp Pro His Ser Lys His Leu Tyr Thr Ala
20 25 30

Asp Met Phe Thr His Gly Ile Gln Ser Ala Ala His Phe Val Met Phe
35 40 45

Phe Ala Pro Trp Cys Gly His Cys Gln Arg Leu Gln Pro Thr Trp Asn
50 55 60

Asp Leu Gly Asp Lys Tyr Asn Ser Met Glu Asp Ala Lys Val Tyr Val
65 70 75 80

Ala	Lys	Val	Asp	Cys	Thr	Ala	His	Ser	Asp	Val	Cys	Ser	Ala	Gln	Gly	
				85					90					95		
Val	Arg	Gly	Tyr	Pro	Thr	Leu	Lys	Leu	Phe	Lys	Pro	Gly	Gln	Glu	Ala	
				100					105					110		
Val	Lys	Tyr	Gln	Gly	Pro	Arg	Asp	Phe	Gln	Thr	Leu	Glu	Asn	Trp	Met	
				115					120					125		
Leu	Gln	Thr	Leu	Asn	Glu	Glu	Pro	Val	Thr	Pro	Glu	Pro	Glu	Val	Glu	
				130					135					140		
Pro	Pro	Ser	Ala	Pro	Glu	Leu	Lys	Gln	Gly	Leu	Tyr	Glu	Leu	Ser	Ala	
				145					150					155		
Ser	Asn	Phe	Glu	Leu	His	Val	Ala	Gln	Gly	Asp	His	Phe	Ile	Lys	Phe	
				165					170					175		
Phe	Ala	Pro	Trp	Cys	Gly	His	Cys	Lys	Ala	Leu	Ala	Pro	Thr	Trp	Glu	
				180					185					190		
Gln	Leu	Ala	Leu	Gly	Leu	Glu	His	Ser	Glu	Thr	Val	Lys	Ile	Gly	Lys	
				195					200					205		
Val	Asp	Cys	Thr	Gln	His	Tyr	Glu	Leu	Cys	Ser	Gly	Asn	Gln	Val	Arg	
				210					215					220		
Gly	Tyr	Pro	Thr	Leu	Leu	Trp	Phe	Arg	Asp	Gly	Lys	Lys	Val	Asp	Gln	
				225					230					235		
Tyr	Lys	Gly	Lys	Arg	Asp	Leu	Glu	Ser	Leu	Arg	Glu	Tyr	Val	Glu	Ser	
				245					250					255		
Gln	Leu	Gln	Arg	Thr	Glu	Thr	Gly	Ala	Thr	Glu	Thr	Val	Thr	Pro	Ser	
				260					265					270		
Glu	Ala	Pro	Val	Leu	Ala	Ala	Glu	Pro	Glu	Ala	Asp	Lys	Gly	Thr	Val	
				275					280					285		
Leu	Ala	Leu	Thr	Glu	Asn	Thr	Phe	Asp	Asp	Thr	Ile	Ala	Glu	Gly	Ile	
				290					295					300		
Thr	Phe	Ile	Lys	Phe	Tyr	Ala	Pro	Trp	Cys	Gly	His	Cys	Lys	Thr	Leu	
				305					310					315		
Ala	Pro	Thr	Trp	Glu	Glu	Leu	Ser	Lys	Lys	Glu	Phe	Pro	Gly	Leu	Ala	
				325					330					335		
Gly	Val	Lys	Ile	Ala	Glu	Val	Asp	Cys	Thr	Ala	Glu	Arg	Asn	Ile	Cys	
				340					345					350		
Ser	Lys	Tyr	Ser	Val	Arg	Gly	Tyr	Pro	Thr	Leu	Leu	Leu	Phe	Arg	Gly	
				355					360					365		

Ala Pro Leu Gln Gly Gln Leu Val Ala Arg Pro Ser Ser Ile Asn Tyr

85										90					95				
Met	Val	Ala	Pro	Val	Thr	Gly	Asn	Asp	Val	Gly	Ile	Arg	Arg	Ala	Glu				
			100					105					110						
Ile	Lys	Gln	Gly	Ile	Arg	Glu	Val	Ile	Leu	Cys	Lys	Asp	Gln	Asp	Gly				
		115					120					125							
Lys	Ile	Gly	Leu	Arg	Leu	Lys	Ser	Ile	Asp	Asn	Gly	Ile	Phe	Val	Gln				
		130				135					140								
Leu	Val	Gln	Ala	Asn	Ser	Pro	Ala	Ser	Leu	Val	Gly	Leu	Arg	Phe	Gly				
145					150					155					160				
Asp	Gln	Val	Leu	Gln	Ile	Asn	Gly	Glu	Asn	Cys	Ala	Gly	Trp	Ser	Ser				
				165					170					175					
Asp	Lys	Ala	His	Lys	Val	Leu	Lys	Gln	Ala	Phe	Gly	Glu	Lys	Ile	Thr				
			180					185					190						
Met	Thr	Ile	Arg	Asp	Arg	Pro	Phe	Glu	Arg	Thr	Ile	Thr	Met	His	Lys				
		195					200					205							
Asp	Ser	Thr	Gly	His	Val	Gly	Phe	Ile	Phe	Lys	Asn	Gly	Lys	Ile	Thr				
		210				215					220								
Ser	Ile	Val	Lys	Asp	Ser	Ser	Ala	Ala	Arg	Asn	Gly	Leu	Leu	Thr	Glu				
225					230					235					240				
His	Asn	Ile	Cys	Glu	Ile	Asn	Gly	Gln	Asn	Val	Ile	Gly	Leu	Lys	Asp				
				245					250					255					
Ser	Gln	Ile	Ala	Asp	Ile	Leu	Ser	Thr	Ser	Gly	Thr	Val	Val	Thr	Ile				
			260					265					270						
Thr	Ile	Met	Pro	Ala	Phe	Ile	Phe	Glu	His	Ile	Ile	Lys	Arg	Met	Ala				
		275					280					285							
Pro	Ser	Ile	Met	Lys	Ser	Leu	Met	Asp	His	Thr	Ile	Pro	Glu	Val					
		290				295					300								

<210> 439

<211> 378

<212> PRT

<213> Homo sapiens

<400> 439

Val	Val	Pro	Ser	Thr	Lys	Asp	Phe	Leu	Val	Gly	Val	Lys	Gly	Ser	Gly
				5					10					15	

Gly	His	Arg	Gly	Gly	Gly	Glu	Met	Ala	Phe	Ser	Gly	Ser	Gln	Ala	Pro
			20					25					30		

0000360"02'29960

Tyr	Leu	Ser	Pro	Ala	Val	Pro	Phe	Ser	Gly	Thr	Ile	Gln	Gly	Gly	Leu
		35					40					45			
Gln	Asp	Gly	Leu	Gln	Ile	Thr	Val	Asn	Gly	Thr	Val	Leu	Ser	Ser	Ser
	50					55					60				
Gly	Thr	Arg	Phe	Ala	Val	Asn	Phe	Gln	Thr	Gly	Phe	Ser	Gly	Asn	Asp
65					70					75					80
Ile	Ala	Phe	His	Phe	Asn	Pro	Arg	Phe	Glu	Asp	Gly	Gly	Tyr	Val	Val
				85					90					95	
Cys	Asn	Thr	Arg	Gln	Asn	Gly	Ser	Trp	Gly	Pro	Glu	Glu	Arg	Lys	Thr
			100					105					110		
His	Met	Pro	Phe	Gln	Lys	Gly	Met	Pro	Phe	Asp	Leu	Cys	Phe	Leu	Val
		115					120					125			
Gln	Ser	Ser	Asp	Phe	Lys	Val	Met	Val	Asn	Gly	Ile	Leu	Phe	Val	Gln
	130					135					140				
Tyr	Phe	His	Arg	Val	Pro	Phe	His	Arg	Val	Asp	Thr	Ile	Ser	Val	Asn
145					150					155					160
Gly	Ser	Val	Gln	Leu	Ser	Tyr	Ile	Ser	Phe	Gln	Asn	Pro	Arg	Thr	Val
				165					170					175	
Pro	Val	Gln	Pro	Ala	Phe	Ser	Thr	Val	Pro	Phe	Ser	Gln	Pro	Val	Cys
			180					185					190		
Phe	Pro	Pro	Arg	Pro	Arg	Gly	Arg	Arg	Gln	Lys	Pro	Pro	Gly	Val	Trp
		195					200					205			
Pro	Ala	Asn	Pro	Ala	Pro	Ile	Thr	Gln	Thr	Val	Ile	His	Thr	Val	Gln
	210					215					220				
Ser	Ala	Pro	Gly	Gln	Met	Phe	Ser	Thr	Pro	Ala	Ile	Pro	Pro	Met	Met
225					230					235					240
Tyr	Pro	His	Pro	Ala	Tyr	Pro	Met	Pro	Phe	Ile	Thr	Thr	Ile	Leu	Gly
				245					250					255	
Gly	Leu	Tyr	Pro	Ser	Lys	Ser	Ile	Leu	Leu	Ser	Gly	Thr	Val	Leu	Pro
			260					265					270		
Ser	Ala	Gln	Arg	Phe	His	Ile	Asn	Leu	Cys	Ser	Gly	Asn	His	Ile	Ala
		275					280					285			
Phe	His	Leu	Asn	Pro	Arg	Phe	Asp	Glu	Asn	Ala	Val	Val	Arg	Asn	Thr
	290					295					300				
Gln	Ile	Asp	Asn	Ser	Trp	Gly	Ser	Glu	Glu	Arg	Ser	Leu	Pro	Arg	Lys
305					310					315					320

Met Pro Phe Val Arg Gly Gln Ser Phe Ser Val Trp Ile Leu Cys Glu
 325 330 335

Ala His Cys Leu Lys Val Ala Val Asp Gly Gln His Leu Phe Glu Tyr
 340 345 350

Tyr His Arg Leu Arg Asn Leu Pro Thr Ile Asn Arg Leu Glu Val Gly
 355 360 365

Gly Asp Ile Gln Leu Thr His Val Gln Thr
 370 375

<210> 440
 <211> 2239
 <212> DNA
 <213> Homo sapiens

<400> 440
 ggaggttgaa gtgagcagag atcatgccag cctgggtgac agtgagactc tgtctcaaac 60
 agaattaagg aaaaaagaaa gaaagaaaaa gagagagagg aaattccagg ccaattgtgg 120
 catagatttt atcatattct ggattttttg gattcttttg tttctcatc actggattca 180
 ggaaagcctg ttgtgtccac catctccaaa ggaggttacc tgcagggaaa tgtaaacggg 240
 aggctgcctt ccctgggcaa caaggagcca cctgggcagg acgccttttc aggaagagac 300
 gccttttcag gaagagacgc cttttcagga agagagaaaag tgcagctgaa gaggaaagtc 360
 actttactga ggggagtcct cattatcatt ggcaccatca ttggagcagg aatcttcata 420
 tctcctaagg gcgtgtctca gaacacgggc agcgtgggca tgtctctgac catctggacg 480
 gtgtgtgggg tctgtgtcact atttggagct ttgtcttatg ctgaattggg aacaactata 540
 aagaaatctg gaggtcatta cacatatatt ttggaagtct ttggtccatt accagctttt 600
 gtacgagtct ggggtggaact cctcataata cgccctgcag ctactgctgt gatatccctg 660
 gcatttgagc gctacattct ggaaccattt tttattcaat gtgaaatccc tgaacttgcg 720
 atcaagctca ttacagctgt gggcataact gtagtgatgg tcctaaatag catgagtgtc 780
 agctggagcg cccggatcca gattttctta accttttgca agctcacagc aattctgata 840
 attatagtcc ctggagttat gcagctaatt aaagggtcaaa cgcagaactt taaagacgcc 900
 ttttcaggaa gagattcaag tattacgcgg ttgccactgg ctttttatta tggaatgtat 960
 gcatatgctg gctgggtttta cctcaacttt gttactgaag aagtagaaaa ccctgaaaaa 1020
 accattcccc ttgcaatatg tatatccatg gccattgtca ccattggcta tgtgctgaca 1080
 aatgtggcct actttacgac cattaatgct gaggagctgc tgctttcaaa tgcagtggca 1140
 gtgacctttt ctgagcggct actgggaaat ttctcattag cagttccgat ctttgttgcc 1200
 ctctctgctt ttggctccat gaacggtggg gtgtttgctg tctccagggt attctatgtt 1260
 gcgtctcgag agggtcacct tccagaaatc ctctccatga ttcatgtccg caagcacact 1320
 cctctaccag ctgttattgt tttgcacct ttgacaatga taatgctctt ctctggagac 1380
 ctgcacagtc ttttgaattt cctcagtttt gccagggtggc tttttatttg gctggcagtt 1440
 gctgggctga tttatcttcg atacaaatgc ccagatatgc atcgctcttt caagggtgcc 1500
 ctgttcattc cagctttgtt ttcttccaca tgctcttcca tgggtgccct ttccctctat 1560
 tcggacccat ttagtacagg gattggtctt gtcactactc tgactggagt ccctgcgtat 1620
 tatctcttta ttatatggga caagaaaccc aggtgggtta gaataatgtc agagaaaata 1680
 accagaacat tacaataaat actggaagtt gtaccagaag aagataagtt atgaactaat 1740
 ggacttgaga tcttggaact ctgcccagg ggagacacaa aatagggatt tttacttcat 1800
 tttctgaaag tctagagaat tacaactttg gtgataaaca aaaggagtca gttattttta 1860
 ttcatatatt ttagcatatt cgaactaatt tctaagaaat ttagttataa ctctatgtag 1920
 ttatagaaag tgaatatgca gttattctat gagtcgcaca attcttgagt ctctgatacc 1980
 tacctattgg ggttaggaga aaagactaga caattactat gtggtcattc tctacaacat 2040

0056770.092000

[illegible]